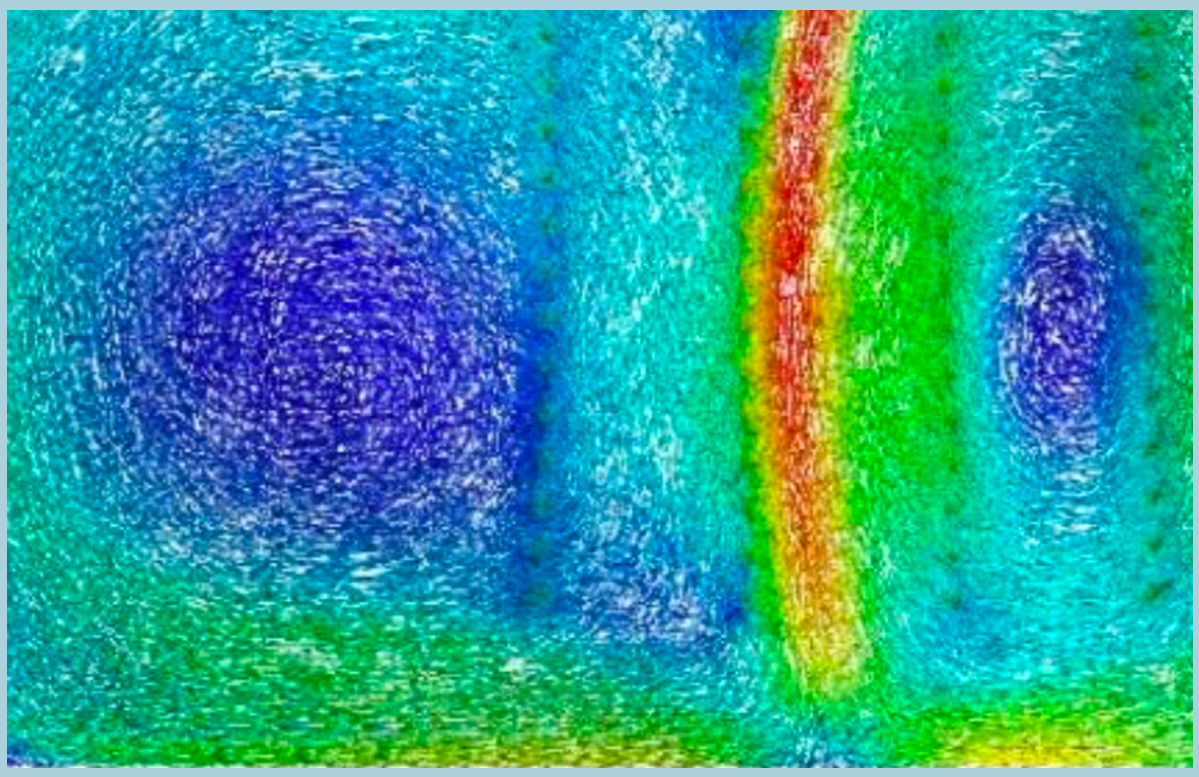
CHEMICAL ENGINEERING RESEARCH

Reports of the 4th year research projects in the Department of Chemical Engineering at Imperial College London

VOLUME 2



Edited by Erich A. Müller

February, 2020

CHEMICAL ENGINEERING RESEARCH

Reports of the 4th year research projects
in the Department of Chemical Engineering
at Imperial College London

Edited by Erich A. Müller

Volume 2

2020

9 781916 005013

© The Author(s) 2020.
Published by Imperial College London, London SW7 2AZ, UK
Contact details – e.muller@imperial.ac.uk

This book is a compilation of manuscripts created as part of a teaching assignment on the 4th year's Chemical Engineering CENG97005 (formerly CE4-01-1) course. Copyright in each paper rests with its authors. The author of each paper appearing in this book is solely responsible for the content thereof; the inclusion of a paper in the book shall not constitute or be deemed to constitute any representation by Imperial College London that the data presented therein are correct or sufficient to support the conclusions reached or that the experiment design or methodology is adequate.

The book is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0). Under this licence, you may copy and redistribute the material in any medium or format on the condition that: you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law. https://creativecommons.org/licenses/by-nc-nd/3.0/.

The information contained in this publication is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall Imperial College London be liable for damages arising from its use.

E-ISBN 978-1-9160050-1-3

First published in 2020

Preface

This volume of *Chemical Engineering Research* collects the unedited research project reports written by 4th year undergraduates (Class of 2020) of the M.Eng. course on Chemical Engineering in the Department of Chemical Engineering at Imperial College London. The research project spans for one term (Autumn) during the last year of the career and has an emphasis on independence, ability to plan and pursue original project work for an extended period, to produce a high quality report, and to present the work to an audience using appropriate visual aids. Students are also expected to produce a literature survey and to place their work in the context of prior art. The papers presented showcase the diversity and depth of some of the research streams in the department, but obviously only touch on a small number of research groups and interests. For a full description of the research at the department, the reader is referred to the departmental website[1].

The papers presented are in no particular order and they are identified by a manuscript number. Some papers refer to appendixes and/or supplementary information which are too lengthy to include. These files are available directly from the supervisors (see supervisor index at the end of the book). Some of the reports are missing, as they have been embargoed waiting for publication in peer-reviewed journals and or patent applications. A few reports correspond to industrial internships, called LINK projects, currently done at Shell.

Cover figure corresponds to a velocity vector profile from a computational fluid dynamics study (taken from the work of Maximos Sifneos and Vinay Patel, manuscript 14).

London, February 2020

[1] https://www.imperial.ac.uk/chemical-engineering

Title Index

paper	Title	page
1	Using Olivine for Carbon Sequestration from Seawater Through Batch Reaction and Ocean Disposal	1
2	Dynamics of Bursting Bubbles	11
3	Towards a More Sustainable Lithium-Ion Battery Future: Recycling LIBs from Electric Vehicles	18
4	Induction of Fatty Acid Synthesis in *Synechocystis* sp. PCC6803 by modification of culturing medium	28
5	The oxidation of Printex-U in NO_2 at various temperatures and concentrations	37
6	A study into the efficacy and durability of eggshell based superhydrophobic coatings formulated under different conditions	47
7	Virus-mimicking Liposomes for Intracellular Drug Delivery	(*)
8	Corrosion-Resistant Sol-Gel Coatings for 304 Stainless-Steel Cookstoves	57
9	In-Depth Systematic Approach for Synthesis of Fe, N Doped Carbon Nanoparticles for ORR Catalysis	67
10	*In Situ* ATR FTIR Spectroscopy of Interactions Between High-Pressure CO_2 and Polymers	77
11	Synthesis and characterisation of hyperporous carbons from benzene-aniline copolymers for CO_2 adsorption	87
12	Prediction of equation of state data of the next-generation hydrofluoroolefin refrigerants with machine learning techniques and artificial neural networks	97
13	Factors Affecting Protein Trans-splicing using Split Inteins	107

paper	Title	page
14	Computational Fluid Dynamics Modelling of a Wood-Fired Cookstove with a Forced Convection Air-Tray	116
15	Performance of Modified Titanate Nanotubes in Adsorptive Desulfurisation Processes	126
16	Global Optimisation and Machine Learning Techniques for Characterising Design Spaces in Pharmaceutical Processes	136
17	Techno-Economic Feasibility Assessment of Oxyfuel Combustion Applied in the Cement Industry	146
18	Modelling the Deactivation of the Shell DeNOx Catalyst	(*)
19	Adsorption Studies in Nanoporous Materials Using the Zero Length Column Technique	158
20	Ceramic Hollow Fibre Supported ZIF-7 Membrane for Hydrogen Separation	168
21	Ultra-Thin Cyclodextrin Membranes with Aligned Intrinsic Pores for Both Polar and Non-Polar Organic Solvent Nanofiltration	177
22	Techno-Economic Assessment of the Direct Air Capture of Carbon Dioxide	187
23	ATR-FTIR Imaging Technique for Study of Factors Influencing CBZ-NIC Cocrystal Tablet Dissolution	197
24	The Development of a High-Pressure Micro-Viscometer	207
25	Machine Learning to Develop an Engineering Correlation for Critical Properties of Fluids	213
26	Electro-oxidation of Glycerol with Pt and Pd Deposition	219
27	Metalated Ion-Exchange Resins as Oxygen Selective Adsorbents for Air Separation	229
28	Dehydrogenation of Formic Acid with Au-Pd Nanoparticles on Titanate Nanotubes	239

paper	Title	page
29	Exploring the stochastic and oscillatory dynamics of distributive double-site phosphorylation systems	(*)
30	A method to calculate interface properties for atomic force microscopy-imaged sessile droplets at a liquid-solid interface	249
31	Red Blood Cell Preservation via Vacuum Desiccation and Storage	(*)
32	Towards Practical Application of Real-Time Optimisation using Gaussian Processes	259
33	Designing polyurethane foams aimed at sportswear market	269
34	A Numerical Investigation into the Role of Viscosity Contrast in the Convective Mixing of Carbon Dioxide and Oil for Geological Carbon Sequestration Applications	278
35	Alcohol-Sensitive Holographic Sensor	288
36	Assessment of solvent performance in CCS process using molecular-based thermodynamic models	297
37	Optimisation and Evaluation of CHP Engine Operational Strategies within Sainsbury's Stores	307
38	Carbon composite anodes in carbonate- and ether-based electrolytes for potassium-ion batteries	317
39	Highly selective microporous polymer membranes for water nanofiltration applications	(*)
40	Effect of high energy facets on the crystal size distributions of seeded batch processes	327
41	Enviro-economic analysis of the production of Dimethyl Ether via Syngas	337
42	Process Domain and Composition Range Reduction Method for Computer-aided Molecular and Process Design	347
43	Exploring the Synthesis of Resin-Based Structured BN and Its Scale-Up Potential	(*)

paper	Title	page
44	Towards the Development of a Computer-Aided Molecular and Process Design Framework for Mixtures as Organic Rankine Cycles Working Fluids	357
45	Evaluation of turbidity and pH technologies for Multi-Functional Micro Sensors (MFMS)	367
46	Modelling of a free-flow photovoltaic-thermal (PV-T) water collector	(*)
47	Taylor bubble characteristics in curved microfluidic channels	377
48	Process modelling: different routes to produce FDCA from lignocellulosic biomass	385
49	Techno-Economic Feasibility Study of Producing CO2 Based Ethylene via Electrochemical Reduction	(*)
50	Investigating Active Learning through Atrivity to Chemical Engineering	395
51	Studying the Micro-Scale Wettability of Multiphase Systems using Atomic Force Microscopy	405
52	Chemical Recycling of PET using Microwave Assisted Catalytic Glycolysis	(*)
53	Investigating the potential of porous ceramic membranes prepared via the adapted freezing technique	415
54	An Environmental and Economic Assessment on the life cycle of Transport fuels: A Study on the Western Australian Heavy Good Vehicles Sector	425
55	Modelling oscillatory flow in helical reactors using adaptive mesh refinement simulations at low Reynolds numbers	434
56	Solution Processed Nanocomposites of Metal-Organic-Frameworks and Microporous Polymers	(*)
57	Hydrogen Infrastructure Spatio-Temporal Optimisation for the UK's Heat Network	444
58	Mechanistic Modelling of RNA Vaccine Production for Quality by Design	454

paper	Title	page
59	pH- and Redox-Responsive Hydrogels in Nanoemulsion Systems for Oral Drug Delivery	464
60	A Two-Step Approach to Modelling Electrodialysis and its Engineering Limitations	(*)
61	Investigating the 4-Micro Channel Design of Solid Oxide Electrolysis Cells for CO2 Electrolysis	(*)

(*) These papers have been removed by request of the authors and/or supervisors

Author and Supervisor index at the end of the book (page 474)

Using Olivine for Carbon Sequestration from Seawater Through Batch Reaction and Ocean Disposal

Arfan Chaudhry and Rhea Patel
Department of Chemical Engineering, Imperial College London, U.K.

Abstract

Global warming has become an increasingly prevalent issue over the past decades, with many recently turning to negative emissions technologies (NETs) to actively reduce the carbon dioxide levels in the atmosphere. One such NET is industrial enhanced weathering, which has gained significant attention as an investment-worthy NET. This report investigates the feasibility of this process, involving the batch reaction of Olivine rocks with seawater, which has been artificially acidified in order to increase its carbonic acid concentration. Other key elements to the process include grinding, mining and transport. Overall, industrial enhanced weathering was found to be infeasible, predominantly due to the unrealistic reactor volume. The key limitation to the scale-up of this process was the slow dissolution rate of Olivine, which manifests itself throughout the results, from the long reaction time, to the large reactor volume and the significant energy requirements for the process. The investigation was extended to compare the batch process to directly adding Olivine to the ocean. This led to a negligible conversion of Olivine by the time it reached the sea floor, and thus was a less favourable alternative to the batch process.
Keywords: *Olivine, carbonic acid, seawater, enhanced weathering, negative emissions technology*

1 Introduction

The current atmospheric level of carbon dioxide (CO_2) is 50% greater than pre-industrial levels (Pires, 2019). These anthropogenic CO_2 emissions have accounted for a 0.2°C increase in the Earth's temperature per decade. By 2017, this had accounted for an overall temperature rise of 1°C beyond pre-industrial levels. This warming has had a considerable negative impact on global ecosystems, in the form of extreme and unprecedented weather conditions (IPCC, 2015). The Paris Agreement of 2015 pledged to keep the average global temperature at a level less than 2°C above pre-industrial levels, whilst actively undertaking efforts to limit this to 1.5°C. The Intergovernmental Panel on Climate Change acknowledged the importance of a 1.5°C target, producing a report on the impacts of allowing global temperatures to rise above this level and the mitigation pathways required to realise this goal (IPCC, 2018).

Negative emissions technologies (NETs) are used to sequester CO_2, reducing the overall CO_2 level in the atmosphere. NETs can either sequester CO_2 directly from the air or indirectly from oceans, as there exists an equilibrium between atmospheric CO_2 and the dissolved CO_2 in seawater, such that a decrease in the CO_2 level in the ocean will, over time, reduce its level in the atmosphere.

This report reviews existing literature concerning NETs, comparing these using a set of key performance indicators (KPIs). These KPIs address cost, annual CO_2 removal potential and resource requirements. The overall objective is to investigate the feasibility of an industrial enhanced weathering process, the idea for which has gained significant attention recently, with certain parties looking to invest in this as a plausible NET. In this process, Olivine, a silicate rock in the Earth's upper mantle, will be used to sequester CO_2 from seawater via a batch reaction. This report will explore the dissolution rates of other minerals and an alternative technique, known as ocean disposal. In this, Olivine is deposited directly into the ocean, reacting with seawater to sequester CO_2.

2 Literature Review

The NETs investigated include those taking place in the ocean, those resembling an industrial process, those involving restoration and those targeting the ground through soil and minerals.

2.1 Ocean

2.1.1 Ocean Fertilisation

Ocean fertilisation (OF) involves stimulating phytoplankton growth by adding nutrients, such as iron, into the ocean. This sequesters CO_2 from the atmosphere. The phytoplankton dies and decomposes, subsequently sinking into the deep ocean, where the carbon is stored (McLaren, 2012). The cost of ocean fertilisation varies from 1-90\$/tCO_2 (Williamson et al., 2012) depending on the nutrients used, with an overall removal potential of 0.91-3.67GtCO_2/yr (Williamson et al., 2012; Harrison, 2017). There are potential side effects of OF, such as an excess of nutrients in the surface ocean, leading to eutrophication. This could result in the development of harmful algal blooms, consequently reducing oxygen levels for aquatic ecosystems. Another issue with OF is the large scale redistribution of nutrients, as a decrease in nutrient supply may negatively impact economic activities, such as fishing (Pires, 2019).

2.1.2 Ocean Liming

Ocean liming involves adding calcium hydroxide or calcium bicarbonate solution to oceans to accelerate the removal of CO_2 from the atmosphere, trapping it in a dissolved inorganic form (McLaren, 2012). The cost of ocean liming is 51-159\$/tCO_2 (McLaren, 2012; Renforth et al., 2013) with a sequestration potential of several gigatones of CO_2 per year (McLaren, 2012). The hydration of lime (slaking) induces large energy requirements for the process of 0.7-6.8GJ/tCO_2 of thermal and electrical energy, and 1.4-1.7 tons of limestone are needed for each ton of CO_2 sequestered (Pires, 2019). An advantage of ocean liming is that it can help to tackle ocean acidification. However, the addition of lime to the oceans creates an imbalance in oceanic ecosystems, the extent to which is unknown (Paquay & Zeebe, 2013).

2.2 Process

2.2.1 Artifical trees

Artificial trees, developed from direct air capture (DAC), is a technology which imitates the actions of biological trees to remove CO_2 directly from the air via a special resin (solid amines). The CO_2 is recovered by hydration of the resin (McGlashan et al., 2012). The cost of artificial trees is expected to vary from 95-300\$/tCO_2 (McGlashan et al, 2012; Fuss et al., 2018) with a removal potential of 0.5-10GtCO_2/yr (Fuss et al, 2018; McLaren, 2012). The energy requirements of the technology are considerably high, at 12.3GJ required per ton of CO_2 removed (Smith et al., 2015). An advantage is that the sorbent can be used repeatedly without degradation (McLaren, 2012). On the other hand, a significant amount of water and electrical energy is required during the hydration process (Workman et al., 2011), and the disposal of CO_2 into geological sinks is an ongoing issue (McGlashan et al., 2012).

2.2.2 Biomass Energy with Carbon Capture and Storage

Biomass Energy with Carbon Capture and Storage (BECCS) refers to the capture and storage of CO_2 emissions from the combustion of biomass. As bioenergy is nominally carbon neutral, the fusion with CCS creates an overall net negative emission (McGlashan et al., 2012; McLaren, 2012). The cost of BECCS ranges from 59-260\$/tCO_2 (McGlashan et al., 2012; McLaren, 2012) and the removal potential lies between 2.4-10GtCO_2/yr (McLaren, 2012). The area required for BECCS is 0.27-0.46ha for every ton of CO_2 removed per year, and the annual water requirement is approximately 720km^3/yr (Smith et al., 2015). The net energy production of each crop is 10.9GJ/tCO_2 (Smith et al., 2015), with a biomass requirement of 0.8Gt for each gigaton of CO_2 sequestered yearly (McGlashan et al., 2012). After the availability of biomass, the scalability of BECCS depends on the development of CCS technology, specifically the creation of a CO_2 piping network with sufficient storage capacity (Azar et al., 2008).

2.2.3 Soda/Lime

The soda/lime process utilises chemical scrubbing with an alkali absorbent to capture CO_2. A conventional scrubbing tower is used to bring the absorbent (aqueous sodium hydroxide) in contact with atmospheric CO_2. Lime ($Ca(OH)_2$) is used to convert the sodium carbonate solution produced back to sodium hydroxide. Production of lime takes place in kilns which can be found in the cement industry. A solution of sodium hydroxide remains which can be recycled back into the scrubbing towers. A precipitate of calcium carbonate is formed in the process (McGlashan et al., 2012). Energy requirements of 2.6–5.3GJ/tCO_2 (Pires, 2019) are required for the input into the kilns and compression of CO_2 for pipeline transportation. The cost of the process varies from 140-600\$/tCO_2 (Pires, 2019; Mclaren, 2012) with a removal potential of 10GtCO_2/yr (McLaren, 2012). This technology has significant electrical and thermal energy demands operating at 900°C due to the calcination step, alongside the use of a kiln requiring fuel to be burnt (most likely, a fossil fuel) (Workman et al., 2011). This would reduce the sequestration potential of this NET.

2.3 Restoration

2.3.1 Afforestation/Reforestation

Reforestation refers to planting trees on land that used to be forest, whereas afforestation refers to this process on areas of land that were not previously forests (Fuss et al., 2018). These are used in combination as a natural NET, with a cost of between 17-100\$/tCO_2 (Smith et al., 2015; McLaren, 2012), and an overall removal potential of 1.5-3.6GtCO_2/year (McLaren, 2012; Fuss et al., 2018). Afforestation/Reforestation (AR) is not a resource-intensive option, with land requirements of 0.29 ha for each yearly ton of CO_2 removed (Pires, 2019), and water requirements of 370km^3/year (Smith et al., 2015). AR is a long-term option for negative emissions, as the S-shaped growth curve of trees limits the uptake potential during the early years (Schlamadinger & Johns, 2007). It also requires strict regulation, as unrestricted afforestation could lead to an increase in food prices of 80% up to 2050, due to competition for land (Kreidenweis et al., 2016). AR has a short storage lifetime (typically 300 years (Lyngfelt, Johansson & Lindeberg, 2019)) compared to less natural NETs. Nevertheless, AR brings along many ecological co-benefits, including better water flows, minimising the potential for flooding and

2

droughts (Schlamadinger & Johns, 2007).

2.3.2 Blue Carbon Restoration

The restoration of coastal ecosystems, such as mangroves and seagrass, can be used to sequester carbon in marine sediments (Greiner et al., 2013). This process, known as blue carbon restoration, has a low cost of 10-20\$/tCO_2 (McLaren, 2012), with an equally low removal potential of 220-773MtCO_2/year (Garrard & Beaumont, 2014; McLaren, 2012). However, the storage lifetime of carbon in these sediments is in the range of ten to thousands of years, due to their anoxic nature, coupled with the fact that they continually accumulate sediment (Greiner et al., 2013). Similarly to AR, blue carbon restoration can generate many environmental benefits, such as improved water quality and nutrient accumulation (Greiner et al., 2013). One key disadvantage is that disturbance of the coastal ecosystem can cause the stored carbon to be released into the atmosphere (Pendleton et al., 2012).

2.4 Soil and Mineral

2.4.1 Biochar

Biochar is a type of charcoal substance, produced by the pyrolysis of biomass, which can be added to soil to enhance its carbon sequestration potential (Mohd et al., 2013). Its cost varies from 30-327\$/tCO_2 (Fuss et al., 2018; Smith, 2016), with a removal potential between 0.9-3GtCO_2/year (McLaren, 2016; Fuss et al., 2016). Advantages of biochar include its long storage lifetime of thousands of years (Lyngfelt, 2019; Mohd et al., 2013), as well as the fact that it can be used on all managed land without drastic changes in land use (Smith, 2016). The production of biochar does not require significant volumes of water or areas for application (<0.27ha per ton of CO_2 removed) (Smith, 2016), although it is expected to need 3.3Gt of biomass for every gigaton of carbon dioxide removed yearly (McGlashan et al., 2012), which is much greater than BECCS. Biochar's ecological benefits include improved soil fertility and reduced leaching of pollutants, but further research is needed to investigate carbon stability in biochar (Mohd et al., 2013).

2.4.2 Soil Carbon Sequestration

Soil carbon sequestration (SCS) involves the natural storage of CO_2 from the air in soil (McLaren, 2012). This process can be enhanced in a range of ways, from changes in tillage practice to the addition of organic amendments. Depending on the method used, SCS has a cost ranging from 0-100\$/tCO_2 (Fuss et al., 2018), amounting in a removal potential of 1.46-5GtCO_2/year (Fuss et al., 2018; Smith, 2016). A key limitation of SCS involves the reversibility of carbon storage in soil (Smith, 2016), the lifetime of which is only 50-80 years (Schmidt et al., 2018). Along with this, land requirements may reach 9ha for each ton of CO_2 removed (Smith, 2016), which is much greater than the application area of biochar.

2.4.3 Enhanced Weathering

Enhanced weathering (EW) involves crushing silicate rocks, such as Olivine, and applying this to large areas of land. The silicate rocks react with the carbonic acid in rainwater, and pass into oceans, forming carbonates (McLaren, 2012). EW incurs a cost of 50-200\$/tCO_2 (Fuss et al., 2018; Strefler et al., 2018) and is capable of removing 0.73-3.67GtCO_2/year (Fuss et al., 2016), with a storage lifetime of over 1000 years (Lenton & Britton, 2006). The application area is expected to be less than 0.0027ha per ton of CO_2 (Smith et al., 2015). The process requires 0.3km^3 of water per year, however the energy required is substantial, with 12.5GJ required per ton of CO_2 removed (Smith et al., 2015), which can mainly be attributed to crushing and grinding. This process is limited by a range of factors, such as the cost of grinding, the mining and transport requirements, as well as the low dissolution rate leading to an overall slow process (Rigopoulos et al., 2018).

3 Methodology

3.1 Enhanced Weathering Process

Olivine (or forsterite, the magnesium-rich Olivine) is a type of ultramafic igneous rock. It is commonly used for enhanced weathering due to its significant carbon sequestration potential and its relatively fast dissolution rate (Moosdorf, Renforth & Hartmann, 2014). Equation 1 details the carbonation reaction of Olivine and carbonic acid to form magnesium carbonate (Griffioen, 2017):

$$Mg_2SiO_4 + 2CO_2 + 2H_2O \rightarrow 2MgCO_3 + H_4SiO_4 \quad (1)$$

Alternative reactions have been proposed, forming magnesium bicarbonate, however these are rarely used (Griffioen, 2017). Figure 1 shows the sequence of units in the process. Upon forming the carbonate precipitate, the rocks can be deposited into the ocean as a slurry.

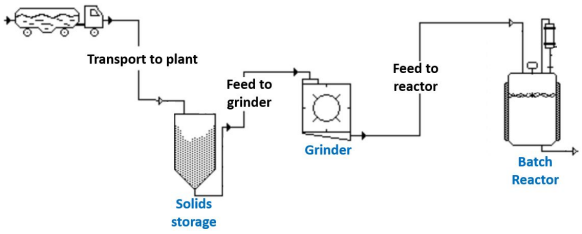


Figure 1: A schematic of the process (created using SuperPro Designer 8.5)

3.2 Reactor Design

The reactor in this process was modelled as an open batch reactor, such that each mole of carbonic acid consumed via the reaction would then

3

be directly replaced by one mole CO_2 from the atmosphere. This leads to an overall constant concentration of carbonic acid in the mixture, although it is simultaneously being consumed by the reaction.

The radius of the reacting Olivine core decreases as the rock reacts. This was accounted for using a shrinking core model under chemical reaction control, where the radius of the core is dependent on time (Equation 2).

$$\frac{r_c}{R} = 1 - \frac{t}{\tau} \text{ where } \tau = \frac{\rho_p R}{2Mr_p r_{As}} \quad (2)$$

r_c is the radius of the reacting core, R is the initial radius, t is time, τ is the reaction time, ρ_p and Mr_p are the $MgCO_3$ density and molar mass respectively, and r_{As} is the Olivine dissolution rate.

The dissolution rate of Olivine (the rate limiting step) was found to be dependent on the temperature (T) and the pH of the reaction mixture (Equation 3, (Rimstidt, Brantley & Olsen, 2012)), where the pH related to the concentration of carbonic acid present (calculations included in Appendix B.1.2).

$$log(r_{As}) = 6.05 - 0.46pH - \frac{3683}{T} \quad (3)$$

As mentioned previously, this overall acid concentration was assumed to be constant, as the batch reactor was open, not sealed, so any carbonic acid removed from the solution was directly replenished by atmospheric CO_2. In order to increase the dissolution rate, the concentration of carbonic acid in seawater was calculated for a scenario in which the seawater was brought to equilibrium with the CO_2 emissions from the effluent of a coal-fired power plant, for which $pCO_2 \approx 0.15$atm, in contrast to the atmospheric concentration of 3.7×10^{-4}atm (Rau & Caldeira, 1999). The calculation for the corresponding carbonic acid concentration can be found in Appendix B.1.1. The dissolution rate was substituted into the mass balance equation for a batch reactor (Equation 4), along with the shrinking core radius as a function of time (Equation 2). This was integrated to obtain Equation 5, which was used to obtain a profile of reactant conversion over time. The full model derivation can be found in Appendix B.1.3.

$$-\frac{dN_R}{dt} = r_{As} \cdot N \cdot 4\pi r_c^2 \quad (4)$$

$$N_{R0}x_R = r_{As} \cdot N \cdot 4\pi R^2 \left[t - \frac{t^2}{\tau} + \frac{t^3}{3\tau^2} \right] \quad (5)$$

N_R is the moles of Olivine, N is the number of particles and x_R is the conversion of Olivine.

From this, the volume of the reactor was determined. This was approximated to the volume of reaction mixture, which was calculated using the stoichiometry of the reaction as well as a set target (calculations included in Appendix B.1.4). It was assumed that each plant had a single batch reactor.

The effect of temperature on the rate of dissolution and overall reactor volume of a single batch was examined by varying the temperature between 298K and 323K. K_{a1} and K_{a2} of carbonic acid are both temperature dependent (equations included in Appendix B.1.5), therefore the rate of dissolution varies with pH and temperature simultaneously. A new rate was calculated for each temperature, and the same reaction model (Equation 5) was then applied.

3.3 Energy Requirements

3.3.1 Reactor Operation

The energy requirement and temperature rise for isothermal (dT/dt) and adiabatic ($q_{exchange} = 0$) operation, respectively, were calculated using the energy balance of a batch reactor (Equation 6, (Leveneur, Vernieres-Hassimi & Salmi, 2016)). The manipulated forms of the energy balance can be found in Appendix B.2.2. The operation was assumed to be isobaric due to the excess of liquid. The enthalpy of reaction was calculated using the enthalpies of formation (Equation 7, data in Appendix B.2.1). Since this energy balance was used to gauge orders of magnitude, the density and specific heat capacity of the reaction mixture were assumed to be that of pure water.

$$V \cdot \rho \cdot C_P \cdot \frac{dT}{dt} - \Delta H_{rxn} \cdot \frac{dN_R}{dt} = q_{exchange} \quad (6)$$

$$\Delta H_{rxn} = \Sigma \Delta H_f^0(products) - \Sigma \Delta H_f^0(reactants) \quad (7)$$

V is the volume of reaction mixture, ρ is the density of the mixture, C_P is the specific heat capacity, ΔH_{rxn} is the enthalpy of reaction, $q_{exchange}$ rate of energy transfer and ΔH_f^0 is the standard enthalpy of formation.

3.3.2 Preheating

In line with the temperatures investigated in Section 3.2, the energy needed to preheat the reaction mixture from ambient conditions (298K) to each of these temperatures was calculated using Equation 8.

$$Q = m \cdot C_p \cdot \Delta T \quad (8)$$

Q is the energy required for preheating and m is the mass of seawater.

This preheating would involve heating the seawater in the batch vessel, prior to adding the Olivine solids. The energy requirement was determined for an individual batch, thus the decrease in batch size required for complete conversion, associated with operating at a higher temperature, was accounted for. Again, the density and specific heat capacity of the reaction mixture were assumed to be that of pure water.

4

3.4 Grinding

The grinding energy requirements were analysed using Rittinger's Law (Equation 9 (Austin, 1973)).

$$E = K_r\left(\frac{1}{d_2} - \frac{1}{d_1}\right) \tag{9}$$

E is the grinding energy per unit mass, K_r is Rittinger's constant, and d_1 and d_2 are the initial and final particle diameters respectively.

Using a literature value of 1.5GJ/t to produce an Olivine particle size of 1μm (d_2) (Renforth, von Strandmann & Henderson, 2015), Rittinger's constant was calculated from the rearranged form of Rittinger's Law, $K_r = \frac{Ed_1d_2}{d_1-d_2}$. A key assumption here was that the Olivine rocks were acquired from the mining site at a radius of 300mm (d_1) (Hangx & Spiers, 2009). This is due to in-pit crushing, a process that takes place at the mine itself.

With Rittinger's constant, d_1 and Equation 9, the grinding energy required to produce a range of particle sizes (d_2) was found, and the final particle size was selected. The grinding energy per batch was calculated by multiplying the grinding energy, E, by the mass of rock required for a single batch. The thermal energy dissipated was calculated as 72% of the input grinding (Bouchard et al., 2017).

3.5 Mining and Transport

Both the mining and transport requirements for the process were analysed with respect to CO_2 emissions, and transport was further investigated to determine the delivery frequency of different vehicles.

3.5.1 Mining

The mining process of Olivine-rich rocks was approximated to that of gold and iron ore mining, as these are all (ultra)mafic rocks and ores with a similar hardness (Giannoulakis, Volkart & Bauer, 2014). Therefore, the CO_2 emission per ton of iron ore mined (7.1 $kgCO_2/t$, (Romao et al., 2012)) was used to find the process' total yearly emissions from mining Olivine.

3.5.2 Transport

Similarly to mining, transport to and from the mine was investigated using literature values for the CO_2 emissions per kilometre per ton of vehicle load (Hartmann et al., 2013):

- Airfreight: $602gCO_2km^{-1}t^{-1}$
- Road heavy goods vehicle: $62gCO_2km^{-1}t^{-1}$
- Diesel rail: $15gCO_2km^{-1}t^{-1}$
- Electric rail: $14gCO_2km^{-1}t^{-1}$
- Inland waterways: $31gCO_2km^{-1}t^{-1}$
- Large tanker: $7gCO_2km^{-1}t^{-1}$

The daily transport frequency was calculated by dividing the yearly Olivine mass requirement by the mass of Olivine carried in one vehicle load. This was limited by the maximum load of the vehicle, rather than the volume capacity, whilst assuming 100% packing density of Olivine particles.

3.6 Further Investigation

3.6.1 Alternative Minerals

The dissolution rates of alternative minerals were investigated, in order assess their sequestration potential compared to Olivine. For each mineral, the dissolution rate dependence on temperature and pH was determined. The minerals investigated were: albite, alunite and serpentine. The full dissolution rates are listed in Appendix B.5.

3.6.2 Ocean Disposal

In order to investigate ocean disposal (where Olivine is deposited directly into the ocean), it was first necessary to find the velocity profile of a single Olivine particle falling through water. For a consistent point of comparison, the radius of the particle was the same as that used in the batch enhanced weathering process. Using equations for the drag force (F_d, Equation 10) and gravitational force (F_g, Equation 11) of this particle, the terminal velocity (u_t) was calculated.

$$F_d = 3\pi d\mu u \tag{10}$$

$$F_g = \frac{4}{3}\pi R^3(\rho_s - \rho_f) \tag{11}$$

μ is the viscosity of seawater, u is the particle velocity, and ρ_s and ρ_f are the rock and seawater densities respectively.

It was necessary to verify the sedimentation regime (which dictates the relevant correlation for F_d), which was done using the particle Reynold's number ($Re_p = \frac{\rho_s d u_t}{\mu}$) From this, the time taken to reach this velocity was determined by plotting Equation 12 (full derivation in Appendix B.6.1).

$$u(t) = \frac{\frac{4}{3}\pi R^3(\rho_s - \rho_f)}{3\pi d\mu u}\left(1 - e^{\frac{-3\pi d\mu t}{m}}\right) \tag{12}$$

Here, m refers to the mass of a single Olivine rock.

For the purpose of the terminal velocity analysis, the radius of the particle was assumed to be R, the initial particle radius, throughout. Although it would be decreasing as it reacts, this would only reduce the terminal velocity, with a negligible impact on the time taken to reach this.

The ocean was then split into batches. Within these, the average pH and temperature was assumed to remain constant, in accordance with an ocean depth pH and temperature profile obtained from literature (data included in Appendix B.6.2). For each batch, the original reaction model (Equation 5) was applied to a single rock falling through

5

the ocean. This could be used to calculate the overall conversion of rock by the time it reaches the sea floor. This process was then scaled up, such that an equivalent mass of rock to that used in the batch process was dumped into the ocean.

The pH of each batch was used to determine the corresponding concentration of carbonic acid, using the equilibrium between dissolved carbon species in the ocean (further details in Appendix B.6.3). This was used to calculate the moles of carbonic acid removed through the process, for comparison with the batch reactor.

4 Results

The annual target for each plant was 5Mt of CO_2 removal, in line with the capacity of similar plants.

4.1 Reactor Design

Figure 2 shows the conversion profile for both reactants, tending to complete conversion over time. Figure 3 shows the negative relationship between reactor volume and reaction temperature (numerical results are included in Appendix B.1.6).

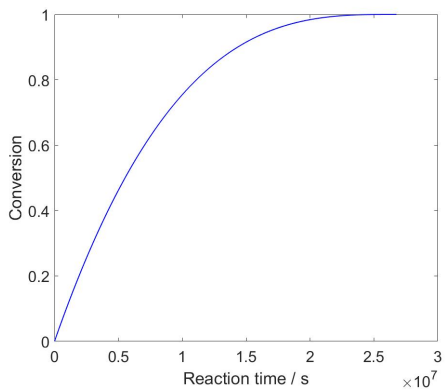


Figure 2: Reactant conversion profile

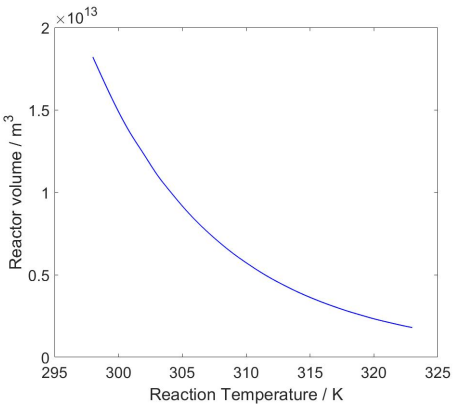


Figure 3: Variation in reactor volume with temperature

A reaction time, τ, of 0.929 years was required to reach complete conversion. For this, a corresponding reactor volume of $1.82\times10^{13}m^3$ was needed. This comprised of only $3\times10^6m^3$ of rock.

4.2 Energy Requirements

For isothermal operation, external cooling was required, between 8.40×10^{14}J (323K) and 8.49×10^{15}J (298K) per batch. However, adiabatic operation led to a 1.1×10^{-4}K temperature rise. Figure 4 shows the energy required to preheat the seawater (prior to the addition of rock). The curve can be attributed to the trade-off between a higher temperature and a reduced mass of seawater per batch (due to a faster rate). Calculations and numerical results are in Appendix B.2.2 and B.2.3.

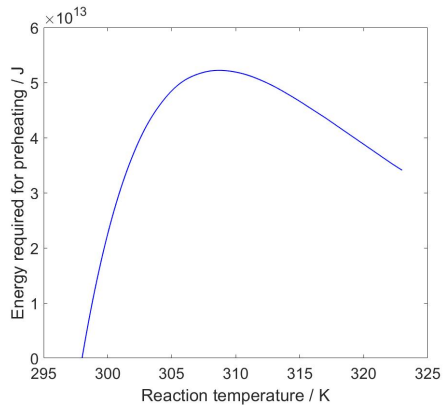


Figure 4: Variation in preheating energy per batch with reaction temperature

4.3 Grinding

Figure 5 shows the steep drop in grinding energy with increasing particle size.

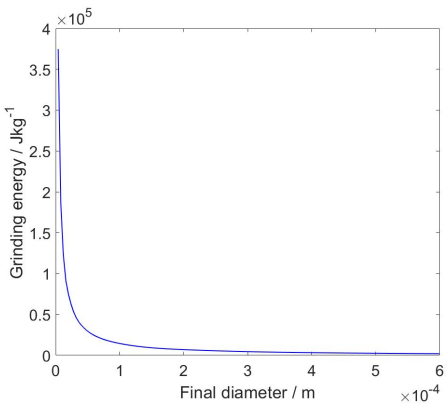


Figure 5: Variation in grinding energy requirements with final particle diameter

From the literature values used, Rittinger's constant was determined to be $1.5Jmkg^{-1}$. A final diameter of 0.02mm was selected, as this best mitigates the trade-off between particle size and grinding energy requirements. Numerical results can be found in Appendix B.3. This gave rise to a grinding energy requirement of 74.5kJ/kg of rock processed. With the yearly removal target in mind, the resulting grinding energy required per batch per batch reached 5.54×10^{14}J, of which 3.99×10^{14}J was dissipated as heat.

4.4 Mining & Transport

4.4.1 CO_2 emissions

The yearly CO_2 emissions from mining amounted to 5.68×10^4t. The yearly CO_2 emissions from

6

transport were: 4810tCO_2km^{-1} for airfreight, 496tCO_2km^{-1} for roads heavy goods vehicles, 112 and 120tCO_2km^{-1} for electric and diesel rail respectively, and 56tCO_2km^{-1} for a large tanker.

4.4.2 Transport frequency

For a road heavy goods vehicle, the volume capacity was 208m^3 (Butcher, 2009), with a load of 520t. This resulted in a transport frequency of 42 trucks per day. A freight train had a volume capacity of 71.3m^3 per wagon, with a maximum load of 60.2t per wagon (Rodrigue, 2019). This gave a transport frequency of 18 trains daily (assuming 20 wagons per train). Finally, for a large tanker, the volume capacity was 915,000m^3, limited by a maximum load of 550,000t (DB Schenker, 2009). This amounted in a transport frequency of one shipment every 25 days.

4.5 Further Investigation

4.5.1 Alternative Minerals

Figures 6 and 7 show how alunite, albite and serpentine's dissolution rates vary with temperature and pH, compared to Olivine.

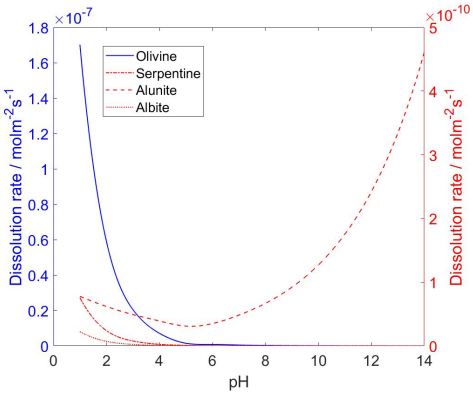


Figure 6: Dependence of dissolution rates on pH

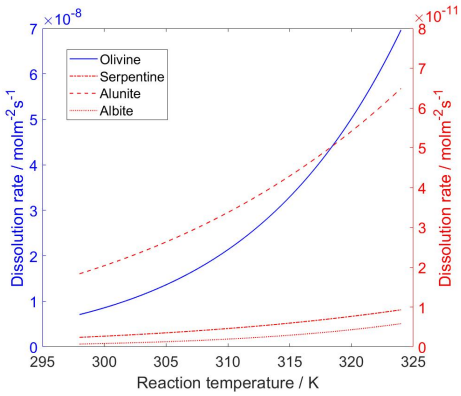


Figure 7: Dependence of dissolution rates on reaction temperature

4.5.2 Ocean Disposal

The terminal velocity of the rock was found to be 3.56×10^{-4}ms^{-1}, reached within 5×10^{-4}s, thus the rock could be assumed to travel at this constant velocity throughout (velocity profile presented in Appendix B.6.1). The rock had a particle Reynold's number of 0.0198. The particle took 0.534 years to fall through the ocean, within which time 2% of the rock converted. When scaled up for the addition of the same mass of rock as the batch process, 2.29×10^9mol of carbonic acid were removed. Appendix B.6.4 contains the numerical results.

5 Discussion

The reaction time for mineral dissolution was extremely long, which can be directly attributed to the slow dissolution rate of Olivine. However, the large reactor volume is the primary indicator of the infeasibility of the process. Since the reaction time was approximately one year, this meant that the yearly target removal had to be achieved in a single batch. The reactor size required for this is infeasible for industrial application. Increasing the dissolution rate by five orders of magnitude would still result in a reactor volume of approximately 10^8m^3, therefore further investigation is needed to find ways to reduce this.

This issue was addressed by operating at a higher reaction temperature. Although this sped up the dissolution rate, heating was highly energy intensive due to the large volume of seawater. The curve also begins to flatten around 320K, therefore increasing the reaction temperature to reduce the reactor volume becomes less beneficial after this point.

One potential solution to the large preheating energy requirements involves harnessing the thermal energy dissipated from the on-site grinding process. The thermal energy dissipated from the grinding of the minerals (per batch) would, in theory, be able to cover the maximum preheating requirement for a single batch, however methods of capturing thermal energy from grinding are currently underdeveloped (Bouchard et al., 2017).

Another potential solution would be to provide the preheating energy from geothermal energy. Geothermal energy is expected to have a global installed capacity of 21.44GW by 2020 (Anderson & Rezaie, 2019), making it an appropriate renewable technology to use alongside NETs.

Of the two reactor operation cases, isothermal operation required a significant amount of cooling, due to the large amount of heat released from the exothermic reaction. However, operating adiabatically showed a negligible temperature rise. This was due to the fact that the thermal energy released by the reaction was distributed across an extremely large volume of seawater. Therefore, the logical choice of operation was adiabatic - since the temperature rise was negligible, it could be likened to isothermal operation without needing any additional energy requirements.

The grinding energy requirements were large, amounting to a similar magnitude as the preheat-

7

ing energy requirements. This is because the slow dissolution rate of Olivine means that a large mass of rock is needed for each batch. However, the yearly grinding energy requirements (also of the order of 10^{14}J) are dependent on the yearly target of the plant, not τ. Therefore, this yearly energy requirement is unrealistic for all batch sizes.

The results for the reactor volume assume that the particles inputted into the reactor have a consistent radius, produced by on-site grinding at the plant. However, grinding would produce a range of particle sizes, following a particle size distribution. Although some may be larger (leading to a slower rate) and others may be smaller (leading to a faster rate), the overall effect of this size variation would not be drastic, thus the reactor volume provides a good indication of the scale required.

The CO_2 emissions from mining did not significantly reduce the net CO_2 sequestration potential of the process, accounting for a 1.14% decrease in potential. Large tankers were found to be the least polluting of the six vehicles investigated, whilst resulting in the most feasible transport frequency (one shipment every 25 days). However, it is important to take into account the lack of flexibility from transport via tankers. The plant would most likely be positioned by the coast (for ease of access to seawater), but if it was positioned further inland for infrastructural reasons, a large tanker would need to be used alongside a form of road or rail vehicle. Therefore, it may be more appropriate to use only rail transport, which was the next least polluting option. This would require a transport frequency of 18 trains per day, whilst producing double the CO_2 emissions of the large tanker, thus this option is also highly infeasible. As a result, the plants would need to be situated directly by the coast, to overcome these difficulties.

The dissolution rates of albite, alunite and serpentine were all approximately three orders of magnitude less than the rate of Olivine. This justifies the initial findings, which highlighted Olivine's relatively fast dissolution rate. The trend of the dissolution rates with temperature were consistent in shape, however the shapes varied significantly when assessing pH dependence. This is due to the fact that certain minerals have a U-shaped dependence on pH (Hellmann, 1994).

The primary downside to ocean disposal of Olivine is that the rock doesn't convert significantly as it falls through the ocean, such that when adding the same amount of rock to the oceans as was added to the batch process, 100 times fewer moles of carbonic acid react. This can be attributed to the dissolution rate being significantly slower, due to the fact that the oceans are relatively alkaline compared to the acidified seawater used in the batch reaction. Since the rock barely converts, it exhibits diminishing returns for every further rock added, as the unreacted rock accumulates at the bottom of the ocean, preventing the layers beneath from reacting. Having said this, dumping Olivine directly into the oceans helps the problem of ocean acidification by reducing the carbonic acid levels, whereas, in the batch process, the seawater has been acidified by contact with the effluent of a power plant, thus disposing this into the ocean actually has a negative impact on the natural pH.

The regulatory framework which aims to limit marine pollution by prohibiting the dumping of waste or other matter into the ocean is encompassed by the London Convention (Akhtar-Khavari, 1997). Amendments should be made to allow for more active manipulation of oceanic chemistry to mitigate climate change. It should also be a pre-requisite for any NET proposing to change ocean chemistry, to fully evaluate the environmental consequences in the context of ocean acidification.

6 Conclusion

The industrial enhanced weathering of Olivine is overall infeasible. The reactor volume exceeds any realistic scale, and the large energy requirements for the process also question the feasibility of operation. However, the CO_2 emissions from mining and transport do not significantly reduce the net CO_2 removal, and the transport frequency is feasible provided that large tankers are used and that the plants are situated at the coast. Disposal of Olivine directly into the ocean is not a feasible alternative, as the conversion of rock as it descends through the ocean is negligible. Thus, despite the flaws with the batch process, it is a better option than ocean disposal. Finally, using alternative minerals does not help the feasibility of the process, as the dissolution rates are lower than that of Olivine.

7 References

Acero, P., Hudson-Edwards, K. & Gale, J. (2015) Influence of pH and temperature on alunite dissolution: Rates, products and insights on mechanisms from atomistic simulation. *Chemical Geology*. 419 1-9.

Akhtar-Khavari, A., (1997) The 1996 Protocol to the 1972 convention on the prevention of marine pollution by dumping of wastes and other matter. *Asia Pacific Journal of Environmental Law*. 2 (2), 201-208.

Anderson, A. & Rezaie, B. (2019) Geothermal technology: Trends and potential role in a sustainable future. *Applied Energy*. 248 18-34.

Argonne National Laboratory, 2019. *Active Thermochemical Tables*. Available at: https://atct.anl.gov/

8

Austin, L. G. (1973) A Commentary on the Kick, Bond and Rittinger Laws of Grinding. *Powder Technology*. 7 315-317.

Azar, C., Lindgren, K., Obersteiner, M. & Riahi, K. (2010). The feasibility of low CO_2 concentration targets and the role of bio-energy with carbon capture and storage (BECCS). *Climatic Change*. 100 (1), 195-202.

Bouchard, J., LeBlanc, G., Levesque, M., Radziszewski, P. & Georges-Filteau, D. (2017) Breaking down energy consumption in industrial grinding mills. *Canadian Mineral Processors Conference*.

Butcher, L. (2009) Lorry sizes and weights. House of Commons.

Charette, M. & Smith, W. (2015) The Volume of Earth's Ocean. *The Official Magazine of the Oceanography Society*.

Chen, Y. & Brantley, S. (1997) Temperature- and pH-dependence of albite dissolution rate at acid pH. *Chemical Geology*. 135 (3-4), 275-290.

Daval, D., Hellmann, R., Martinez, I., Gangloff, S. & Guyot, F. (2013) Lizardite serpentine dissolution kinetics as a function of pH and temperature, including effects of elevated pCO_2. *Chemical Geology*. 351 245-256.

DB Schenker. (2009) Our Freight Wagons.

Fuss, S., Jones, D., Kraxner, F., Peters, G. P., & Smith, P. (2016) Research priorities for negative emissions. *Environmental Research Letters*. 11 (11).

Fuss, S., Lamb, W., Callaghan, M., Hilaire, J. & Creutzig, F. (2018) Negative emissions—Part 2: Costs, potentials and side effects. *Environmental Research Letters.* 13 (6).

Garrard, S. & Beaumont, N. (2014) The effect of ocean acidification on carbon storage and sequestration in seagrass beds; a global and UK context. *Marine Pollution Bulletin*. 86 (1-2).

Giannoulakis, S., Volkart, K. & Bauer, C. (2014) Life cycle and cost assessment of mineral carbonation for carbon capture and storage in European power generation. *International Journal of Greenhouse Gas Control*. 21 140-157.

Greiner, J., McGlathery, K., Gunnell, J. & McKee, B. (2013) Seagrass Restoration Enhances "Blue Carbon" Sequestration in Coastal Waters. *PLoS ONE*.

Griffioen, J. (2017) Enhanced weathering of Olivine in seawater: The efficiency as revealed by thermodynamic scenario analysis. *Science of the Total Environment*. 575 536-544.

Hangx, S. J. T. & Spiers, C. J. (2009) Coastal spreading of Olivine to control atmospheric CO_2 concentrations: A critical analysis of viability. *International Journal of Greenhouse Gas Control.* 3 (6), 757-767.

Harrison, D. (2017) Global negative emissions capacity of ocean macronutrient fertilization. *Environmental Research Letters.*

Hartmann, J., West, A., Renforth, P., Kohler, P., De La Rocha, C., Wolf-Gladrow, D., Durr, H. & Scheffran, J. (2013) Enhanced chemical weathering as a geoengineering strategy to reduce atmospheric carbon dioxide, supply nutrients, and mitigate ocean acidification. *Reviews of Geophysics*. 51 (2), 113-149.

Hellmann, R. (1994) The albite-water system: Part I. The kinetics of dissolution as a function of pH at 100, 200 and 300°C. *Geochimica et Cosmochimica Acta*. 58 (2), 595-611.

Houghton, R. A. (2003) The Contemporary Carbon Cycle. In: Holland, H. & Turekian, K. (eds.) *Treatise on Geochemistry*. Elsevier. pp. 473-513.

Intergovernmental Panel on Climate Change. (2018) *Special Report: Global Warming of 1.5^0C*.

Kleypas, J., Feely, R., Fabry, V., Langdon, C., Chris, S., & Robbins, L. (2006). Impacts of Ocean Acidification on Coral Reefs and Other Marine Calcifiers: A Guide for Future Research. *Report of a Workshop Held*. 18.

Kreidenweis, U., Humpenöder, F., Stevanović, M., Bodirsky, B., Kriegler, E., Lotze-Campen, H. & Popp, A. (2016) Afforestation to mitigate climate change: impacts on food prices under consideration of albedo effects. *Environmental Research Letters.*

Lampitt, R., Achterberg, E., Anderson, T. & Hughes, J. (2008). Ocean fertilization: a potential means of geoengineering?. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 366 (1882), 3919-3945.

Lenton, T. & Britton, C. (2006) Enhanced carbonate and silicate weathering accelerates recovery from fossil fuel CO_2 perturbations. *Global Biogeochemical Cycles*.

Leveneur, S., Vernieres-Hassimi, L. & Salmi, T. (2016) Mass energy balances coupling in chemical reactors for a better understanding of thermal safety. *Education for Chemical Engineers*. 16 17-28.

Lyngfelt, A., Johansson, D. & Lindeberg, E. (2019) Negative CO_2 emissions - An analysis of the retention times required with respect to possible carbon leakage. *International Journal of Greenhouse Gas Control*. 87 27-33.

McGlashan, N., Shah, N., Caldecott, B. & Workman, M. (2012) High-level techno-economic assessment of negative emissions technologies. *Process Safety and Environmental Protection.* 90 (6).

McLaren, D. (2012) A comparative global assessment of potential negative emissions technologies. *Process Safety and Environmental Protection*. 90 (6), 489-500.

9

Millero, F., Pierrot, D., Lee, K., Wanninkhof, R., Feely, R., Sabine, C., Key, R. & Takahashi, T. (2002) Dissociation constants for carbonic acid determined from field measurements. *Deep-Sea Research*. 1705-1723.

Mohd, A., Ghani, W. A., Resitanim, N. Z. & Sanyang, L. (2013) A Review: Carbon Dioxide Capture: Biomass-Derived-Biochar and Its Applications. *Journal of Dispersion Science and Technology*. 34 (7), 974-984.

Mondal, B., Ghosh, D. & Das, A. (2009) Thermochemistry for silicic acid formation reaction: Prediction of new reaction pathway. *Chemical Physics Letters*. 478 (4-6), 115-119.

Moosdorf, N., Renforth, P. & Hartmann, J. (2014) Carbon Dioxide Efficiency of Terrestrial Enhanced Weathering. *Environmental Science & Technology.* 48 (9), 4809-4816.

NIST, 2018. *NIST Chemistry WebBook*. Available at: https://webbook.nist.gov/chemistry/

Paquay, F. & Zeebe, R. (2013). Assessing possible consequences of ocean liming on ocean pH, atmospheric CO_2 concentration and associated costs.*International Journal of Greenhouse Gas Control*. 17, 183-188.

Pendleton, L., Donato, D., Murray, B. & Crooks, S. (2012) Estimating Global "Blue Carbon" Emissions from Conversion and Degradation of Vegetated Coastal Ecosystems. *PLoS ONE*. 7 (9).

Pires, J. C. M. (2019) Negative emissions technologies: A complementary solution for climate change mitigation. *Science of the Total Environment*. 672 502-514.

Rau, G. & Caldeira, K. (1999) Enhanced carbonate dissolution: a means of sequestering waste CO_2 as ocean bicarbonate. *Energy Conversion Management.* 40 (17), 180-1813.

Renforth, P., Jenkins, B. & Kruger, T. (2013) Engineering challenges of ocean liming. Energy. 60 442-452.

Renforth, P., von Strandmann, P. A. E & Henderson, G. M. (2015) The dissolution of Olivine added to soil: Implications for enhanced weathering. *Applied Geochemistry*. 61 109-118.

Rigopoulos, I., Harrison, A. L., Delimitis, A., & Ioannou, I. (2018) Carbon sequestration via enhanced weathering of peridotites and basalts in seawater. *Applied Geochemistry*. 91 197-207.

Rodrigue, J. (2019) Tanker Sizes and Classes: The Geography of Transport Systems.

Romao, I., Nduagu, E., Fagerlund, J., Gando-Ferreira, L. & Zevenhoven, R. (2012) CO_2 fixation using magnesium silicate minerals. Part 2: Energy efficiency and integration with iron-and steelmaking. *Energy*. 41 (1), 203-211.

Rimstidt, J., Brantley, S., & Olsen, A. (2012) Systematic review of forsterite dissolution rate data. *Geochimica et Cosmochimica Acta*. 99 159-178.

Schlamadinger, B. & Johns, T. (2007) Afforestation, Reforestation and Reduced Deforestation to Sequester Carbon and Reduce Emissions. In: Freer-Smith, P., Broadmeadow, M. & Lynch, J. (eds.). *Forestry and Climate Change*. Oxfordshire, CABI. pp. 73-79.

Schmidt, H., Anca-Couce, A., Hagemann, N., Werner, C., Gerten, D., Lucht, W. & Kammann, C. (2018) Pyrogenic carbon capture and storage. *Global Change Biology Bioenergy.*

Sijp, W. & England, M. (2015). The effect of low ancient greenhouse climate temperature gradients on the ocean's overturning circulation. *Climate of the Past Discussions*. 11 4787-4810.

Smith, P. (2016) Soil carbon sequestration and biochar as negative emission technologies. *Global Change Biology*. 22 (3).

Smith, P., Davis, S., Creutzig, F., Fuss, S. & Mix, J. (2015) Biophysical and economic limits to negative CO_2 emissions. *Nature Climate Change.*

Soli, A & Byrne, R. (2002) CO_2 system hydration and dehydration kinetics and the equilibrium CO_2/H_2CO_3 ratio in aqueous NaCl solution. *Marine Chemistry*. 78 (2-3), 65-73.

Strefler, J., Amann, T., Bauer, N., Kriegler, E. & Hartmann, J. (2018) Potential and costs of carbon dioxide removal by enhanced weathering of rocks. *Environmental Research Letters*. 13.

United Nations. (2015) *Paris Agreement*.

Williamson, P., Wallace, D., Law, C., & Boyd, P. (2012). Ocean fertilization for geoengineering: A review of effectiveness, environmental impacts and emerging governance. *Process Safety and Environmental Protection*. 90 (6), 475-488.

Workman, M., McGlashan, N., Chalmers, H. & Shah, N. (2011). An assessment of options for CO_2 removal from the atmosphere. *Energy Procedia*. 4, 2877-2884.

10

Dynamics of Bursting Bubbles

Keneu Kuek
Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom

Bursting bubbles is the main source of aerosols from champagne which ultimately compliments the dining experience. The release of aromatic compounds is via evaporation and the mass transfer of aromatic compounds from the liquid phase into the atmosphere. The motivation of this study is to elucidate the factors governing the ejection of droplets into the surroundings. By performing numerical simulations on a Computational Fluid Dynamics (CFD) software, this study investigates the dynamics of a bubble bursting in three different Laplace numbers $La = 2 \times 10^3, 2 \times 10^4, 2 \times 10^5$ by varying the liquid viscosity. The collapse of the bubble cavity at the liquid surface upon bursting leads to the ejection of a liquid jet which breaks up into droplets. The characteristics of jet droplets (e.g. number, sizes, velocity) released were analysed to determine the robustness of the droplet ejections. A lower viscosity gives a more robust droplet ejection, which ultimately promotes evaporation and mass transfer.

I. BACKGROUND

Wherever water is present, there is bound to be the formation of bubbles on the surface. From oceans, hot springs, pools, streams, fountains, food and drinks, bubbles are ubiquitous.

A single bubble bursting in the ocean can have zero effect on the ecosystem. However, a very large magnitude of bubbles bursting ($\sim 10^{18}$ to 10^{20} per second) in the ocean gives rise to ocean spray, a dominant player in the Earth's climate system.[1]

In the ecosystem, bursting bubbles play a key role in the transmission of chemicals and biological substances between the air and liquid phase. Compositions (e.g. viruses, bacteria, toxins) within contaminated bubbles are released upon bursting via sneezing or naturally-occurring sources in the environment.[2] [3] [4] [5]

In the F&B industry, the dynamics of the bubble bursting phenomenon can serve desirable and has been a subject of interest.[6] [7] In fizzy drinks and champagne, this phenomenon has olfactory and gustatory roles by releasing aromatic compounds thereby enhancing the consumers' experience.[8] [9] [10] [11]

Therefore, it can be seen that this humble phenomenon is not to be overlooked as it proves significant to major concerns and revolutionary aspects from climate change, industry, epidemics and public health.

The study on bursting bubbles started in 1953 by Woodcock et al[12] has come a long way to the current access to wide documentation of numerical and experimental studies.([13], [14], [6], [15], [16])

When a bubble reaches the free surface, the thin liquid film partitioning the bubble and the atmosphere ruptures leading to film drops.[17] An unstable opened cavity is then produced which eventually collapse [18] [19] [20] leading to the formation of a rising central jet, catapulting jet drops vertically above the surface.[21] Pressure gradient forces the gas out to the atmosphere. Greater pressure outside pushes on the bubble interface causing it to deform, producing a jet. Another important factor is the surface free energy. The deformation caused by surface tension minimises the surface energy.

II. PROBLEM STATEMENT AND FORMULATION

A. Governing Equations

1. Flow Equations

The pressure and velocity fields are described by the Navier-Stokes equations for incompressible Newtonian fluids. The continuity equation and the momentum balance are as follows:

$$\nabla \cdot \mathbf{u} = 0 \ , \tag{1}$$

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) = -\nabla P + \rho \mathbf{g} + \mu \nabla \cdot (\nabla \mathbf{u} + \nabla \mathbf{u}^T) + \int_A \sigma \kappa \mathbf{n} dA, \tag{2}$$

where $\mathbf{u}$, P, g, ρ, μ, σ and κ denotes velocity, pressure, gravitational acceleration, density, viscosity, surface tension and the mean curvature of the interface respectively. The final term of the right-hand side of equation 2 is the surface tension force.

Applying the key concept of dynamic similarity, non-dimensionalisation had been employed on the equations above to reduce the number of parameters and simplify the analysis. The initial bubble radius, R_o and the capillary breakup time, $t_c = \sqrt{\rho {R_o}^3/\sigma}$ are selected as the length and time scales respectively. Upon applying the scales, three dimensionless numbers arose: Reynolds $Re = \rho R_o u/\mu$, Weber $We = \rho R_o u^2/\sigma$ and Froude $Fr = u/\sqrt{gR_o}$. The way the characteristic velocity

$\widetilde{u} = R_o/t_c$ is formulated enables the Weber number to equate to 1, Re to be related to the Ohnesorge number, $Re = Oh^{-1}$ and the Froude number to the Bond number $1/Fr^2 = Bo$. Therefore, the dimensionless momentum equation is as follows:

$$\tilde{\rho}\left(\frac{\partial \tilde{\boldsymbol{u}}}{\partial \tilde{t}} + \tilde{\boldsymbol{u}} \cdot \tilde{\nabla}\tilde{\boldsymbol{u}}\right) + \tilde{\nabla}\tilde{P} = Bo\,\tilde{\rho}\mathbf{i}_z + Oh\,\tilde{\nabla}\cdot(\tilde{\nabla}\tilde{\boldsymbol{u}} + \tilde{\nabla}\tilde{\boldsymbol{u}}^T) + \int_A \tilde{\sigma}\kappa\mathbf{n}dA, \quad (3)$$

Using a Computational Fluid Dynamics (CFD) software, the numerical simulation was performed by using the Level Contour Reconstruction Method, a hybrid method which combines the characteristics of the Front Tracking and Level Set methods. Full details of the numerical solution method for the velocity, pressure and interface dynamics are presented in several articles of Shin et al. [22] [23] [24] [25] [26]

B. Numerical Configuration

A numerical hole was applied on the thin film which partitions the bubble at rest on a gas-liquid interface from the atmosphere to initialise the simulation.

For velocity, the Neumman's boundary condition was applied on the entire box except for the bottom boundary at which the Dirichlet boundary condition was applied. As for pressure, the Neumman's boundary condition was applied on the bottom and sides of the box while the Dirichlet's boundary condition was applied on the top face of the box.

The size of the calculation domain is $6R_o \times 6R_o \times 12R_o$. For parallelisation, the domain is divided into $4 \times 4 \times 4 = 64$ subdomains. Each subdomain holds a regular grid mesh of $32 \times 32 \times 64$ cells. The global resolution in the entire domain is then $128 \times 128 \times 256$.

This study investigates the effect of viscosity on the dynamic behaviour of the bubble bursting phenomenon. Viscosity is varied with the La number. According to equation 3, having set $\tilde{\sigma}$, $\tilde{\rho}$ and R_o constant at unity for simulation purposes and simplification, the Oh number, given by $Oh = \mu/\sqrt{\rho\sigma R_o}$, will be the dimensionless parameter which represents viscosity of the liquid, $\mu_l = Oh$. This then sets viscosity of air, $\mu_g = 1/55$ from the relationship $\mu_g = \mu_l/55$. Noting the definition of $La = \sigma\rho R_o/\mu^2$, La will effectively become $La = 1/\mu^2$. Therefore, the following relationship holds, $\mu_l = Oh = 1/\sqrt{La}$, hence allowing the La parameter to be used to impose different liquid viscosity on the system. The Bo number, $Bo = \rho g R_o/\sigma$ postulates the relative effects of gravity compared to surface tension. Setting Bo to 10^{-3}, a number sufficiently small allows gravitational effects to be neglected in this numerical study. It also portrays that surface forces dominate over gravitational forces which ultimately gives a perfectly spherical bubble shape. Three La numbers of magnitude 2×10^3, 2×10^4 and 2×10^5 were selected for this numerical study.

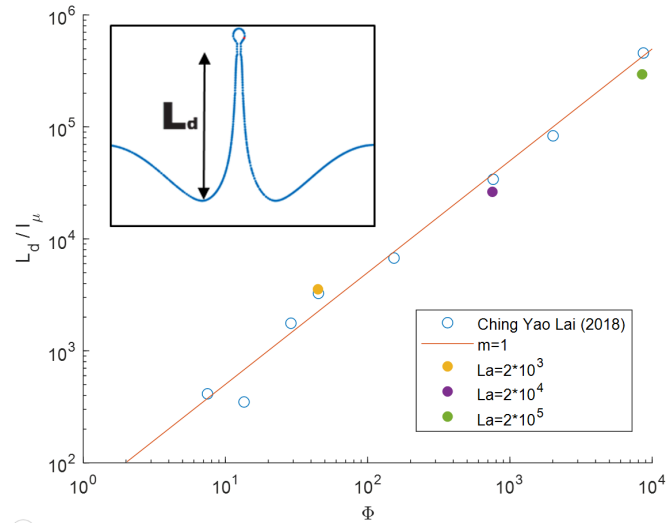


FIG. 1. Comparison between the numerical results (scattered) and the scaling argument(solid line)

C. Method Validation

The method used to validate the numerical results generated was based on whether their trend obeys the scaling laws obtained by Ganan-Calvo.[30] The robustness of Ganan-Calvo's scaling law to characterise the jet variables at the first droplet release was verified by Ching-Yao Lai et al.[16] The relationships involving the dimensionless parameter $\phi \equiv \sqrt{La}(\sqrt{La/La^*} - 1)$ and L_d/l_μ was plotted. $La^* = 500$, is the critical La number in which bubbles are only observed above that value. L_d represents the length of the jet at the instance of time in which the first droplet is formed and about to be released (as illustrated in figure 1) whereas $l_\mu \equiv \mu^2/\rho\sigma$ which can be further simplified to $l_\mu = 1/La$. The numerical results obtained which follows the exact trend as the results from Ching-Yao Lai et al [16] indicates a good agreement has been achieved between the numerical results in this study and the scaling laws, thus validating the results.

III. RESULTS AND DISCUSSION

A. Temporal Evolution of Interface

Figure 2 describes the temporal evolution of the bursting bubble at the intermediate La, 2×10^4. At $t = 0$, the nucleation of a hole at the thin film which partitions a bubble at rest on a gas-liquid interface from the atmosphere occurred. Consequently, as shown at $t = 0.10$, the nucleation of the hole led to a subsequent collapse of the gas cavity which was instigated by the low capillary pressures which appeared suddenly around the base of

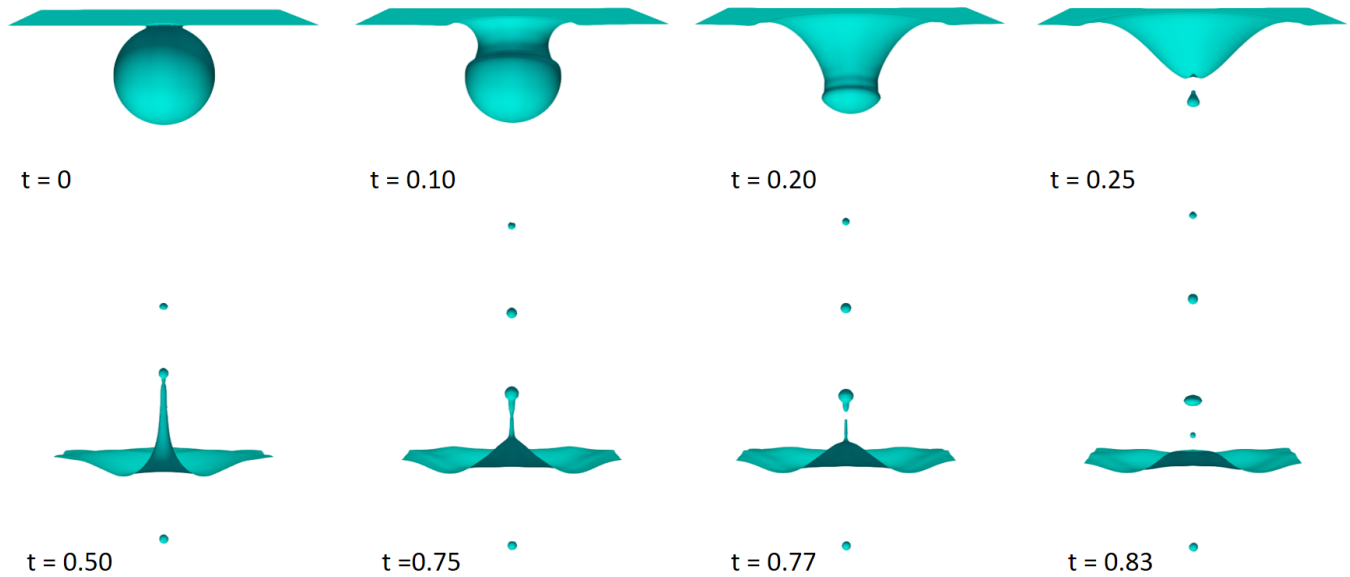


FIG. 2. Temporal evolution of the bursting bubble at $La = 2 \times 10^4$

the cavity's film. The widening of the hole induced by capillarity, generated capillary waves which propagated along the bubble interface.[19] [20] [18] [27]

Simultaneously, pressure gradients and capillarity caused the further deformation of the bubble's interface whereby radial velocity converges to the centre as depicted by Gordillo et al[29]. The crests of the capillary waves merged before they reach the bottom vertex. This resulted in the shape of a truncated cone as observed at $t = 0.25$.

Once the capillary waves reached the nadir (the lowest point of the cavity), the ejection of a Worthington-like jet[28] occurred as depicted at $t = 0.50$.

A relationship exists between the wavelength of the capillary waves and the jet velocity, V_{jet} as studied by Gordillo et al.[29] The relationship has 3 modes depending on the critical Oh, Oh_c, which has a value of 0.02: Below Oh_c, at Oh_c and above Oh_c. Below Oh_c, wavelength, λ^* increases with Oh (noting that Oh equates to μ_l) as $\lambda^* \propto Oh^{1/2}$. Due to the small magnitude of Oh, the relationship $V_{jet} \propto Oh^{1/2}[1 - (Oh/Oh_c)^{1/2}]^{-3/2}$ can be reduced to $V_{jet} \propto Oh^{1/2}$. V_{jet} increases with viscosity until $Oh = Oh_c$.

Qualitative observations of this relationship can be made on capillary waves with Figure 3. The relationship holds in this study because as La becomes bigger, the viscosity becomes smaller which causes capillary wavelengths to decrease and hence, jet velocity decreases. Based on Figure 3, the earlier ejection of the jet indicates a higher jet velocity and vice versa.

Based on Newton's Third Law, the total momentum is conserved via the downward release of a bubble at the very instant a jet emerges at t=0.25 in which both bubble and jet are of equal and opposite momentum. The jet

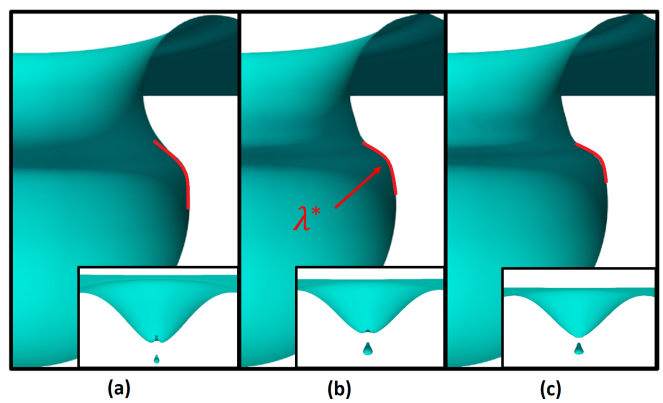


FIG. 3. Bubble Collapse at t=0.10 with the Resulting Evolution at t=0.25 (a) $La = 2 \times 10^3$ (b) $La = 2 \times 10^4$ (c) $La = 2 \times 10^5$

behaviour as time progresses is mapped out on $t = 0.50$, $t = 0.75$, $t = 0.77$ and $t = 0.83$. The phenomenon in which the jet releases and breaks into droplets is a result of Plateau-Rayleigh instability. This occurs due to discontinuities in pressure portrayed by a high pressure region sandwiched between two low pressure regions within the column of the jet as described in the Young Laplace equation:

$$\Delta P = \sigma(1/R_1 + 1/R_2), \tag{4}$$

where ΔP, R_1 and R_2 represent the increase in pressure, the radius of the column and the radius of the curvature of perturbation respectively. Under constant surface tension, column curvature affects internal pressure in which smaller radius at a pinch point results in relatively high pressure within that region compared to the neighbouring regions with the wider column radius.

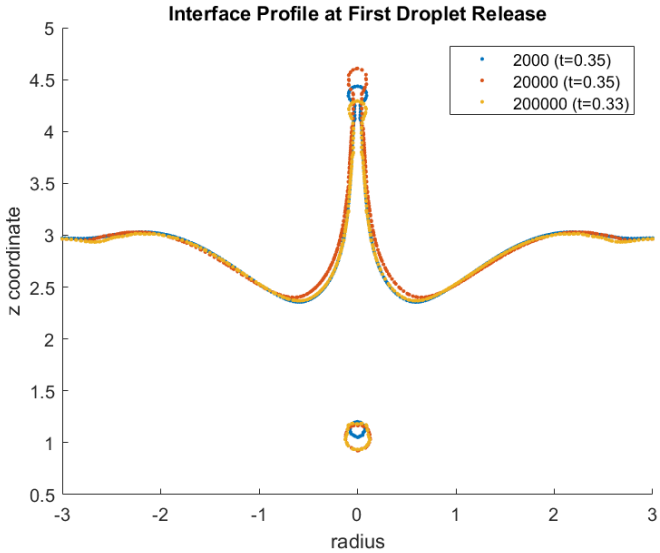


FIG. 4. Interface Profile at First Droplet Release used to obtain Jet Lengths (a) $La = 2 \times 10^3$ (b) $La = 2 \times 10^4$ (c) $La = 2 \times 10^5$

Between t=0.77 and t=0.83, the coalescence of the two lowest droplets was effectively caused by capillarity.

B. Jet Lengths

As studied on Figure 4, the L_d obtained in La 2×10^3, 2×10^4, 2×10^5 were 1.900, 1.913 and 1.710 respectively, concluding the jet length increases in the order of La 2×10^5, 2×10^3, 2×10^4. Solely considering the effects of viscosity which resist deformation, the jet length should increase if viscosity decreases as shown in La 2×10^4 as compared to La 2×10^3. However, there is a significant drop in jet length from La 2×10^4 to La 2×10^5. With this scenario of perturbations taking place and the fact that viscosity only acts to slow down the growth or decay of a given perturbation within the fluid, the shortest jet length observed in La 2×10^5 can be explained by the viscosity being sufficiently low allowing stronger propagation of perturbations. An accompanying scenario in this case involves the surface forces being dominant over the viscous forces as depicted by the high La number. These two factors led to the Plateau-Rayleigh instability occurring faster. Hence, the droplet is released earliest among the other La's. This trend justifies the presence of a trade-off between jet length and early droplet release stage as viscosity is varied.

C. Droplets Number, Size, Velocity and Height

Both qualitative and quantitative observations can be made on the number, sizes and velocity of droplets across the three La's (see figure 5), which has been highly dependent on the viscosity. Higher La means lower viscosity of the liquid. Therefore, liquid with higher La has lower resistance to deformation and hence, indicated a higher magnitude of velocities. In terms of interface shape, a higher magnitude of velocities give rise to perturbations and thus, the curvature of the interface.

At t = 0.10, velocity profile increases from $La = 2 \times 10^3$ to $La = 2 \times 10^5$ hence justifying the initial jet velocities as discussed in the temporal evolution. Higher velocity contributes to a higher frequency of the capillary waves and thus, a smaller capillary wavelength, which in turn gives a lower jet velocity, hence concurring with Gordillo et al.[29]

Number, sizes and velocity of bubble and droplets are studied throughout t=0.50, t=0.70 and t=0.80.

Effects of viscosity can be observed in the jet breakups via the increasing trend on the number of jet droplets at the increment of 1 with the increments of La. For instance, La 2×10^3, 2×10^4 and 2×10^5 produced a total number of droplets of 3, 4 and 5 respectively. This implies that viscosity plays an important role in the breakup of the jet column, by having a damping effect on the perturbations as a consequence of the bubble burst. Detailed studies into the damping effect of viscosity on perturbations such as capillary waves were conducted by U.S. Jeng et al [31] and this effect was acknowledged in Ghabache et al [18] and Krishan et al [27]. Higher viscosity gives a more substantial damping effect. Figure 6 reveals the pressure maps at the critical points of time after the primary and secondary breakups of the jet column to be studied. Figure 6 (a) and (b) provide insights into the size and coalescence of the third droplet upon tertiary breakup. At t=0.77, no breakups of the column occurred despite having observed high curvature (as discussed previously with equation 4). This proves that viscosity has an opposing effect against the Plateau-Rayleigh instability in the breakups of liquid columns. In this case, it is explicit that viscosity has manifested a more dominant effect over the perturbations. The damping effect on perturbations is sufficiently high which allows the coalescence to occur before the breakup of the liquid column. The smallest radius at the bottom end of the tertiary droplet resulted in the highest relative pressure within the region. This effectively forces the liquid within the region to move towards the top end, down the pressure gradient. Effectively, as shown in figure 4b at t=0.80, this leads to a successful coalescence of the tertiary droplet.

Symptoms of the Plateau-Rayleigh instability become more conspicuous with larger La. More droplets are formed from the jet column rupture because the effects of the perturbations become more significant as viscosity decreases, as shown in figure 6 (c) and (d). A greater degree of curvature gives a smaller radius and hence, a greater increase in pressure within the liquid column. Ultimately, this results in breakup of liquid column into droplets.

Across all the La's, it is observed that the droplets produced shrink as time progresses. This can be explained by the effects of surface tension. Surface ten-

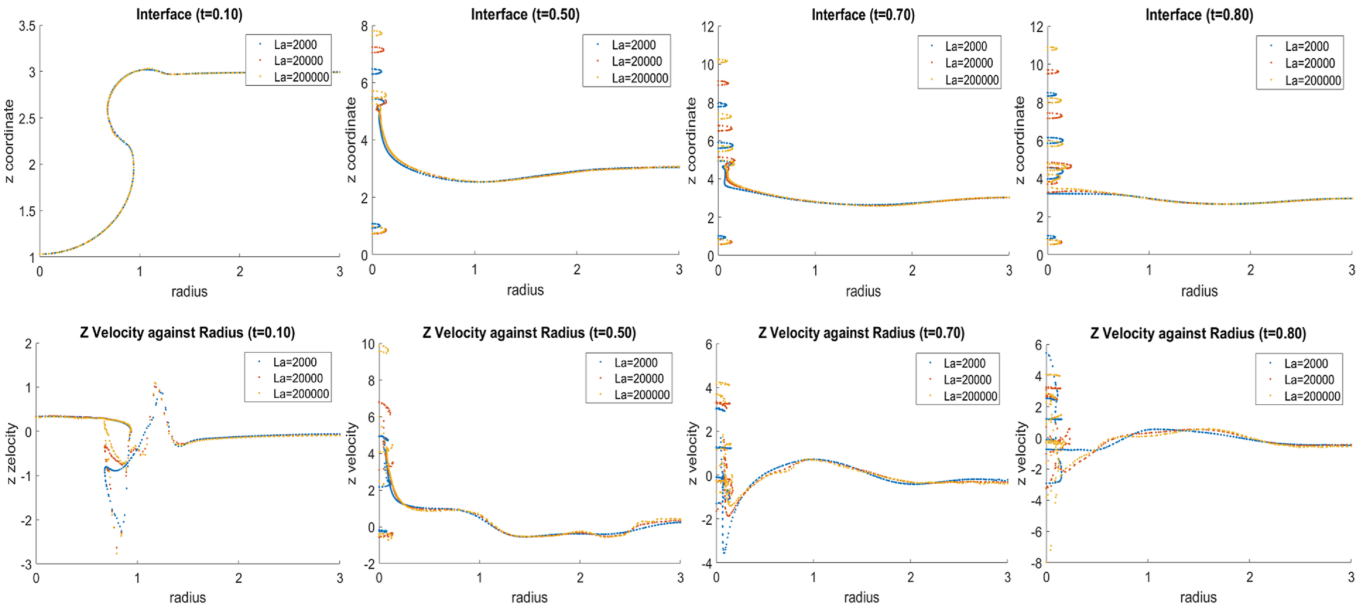


FIG. 5. Half plots of the interface with their corresponding Z velocity profiles at different times (due to symmetrical profiles)

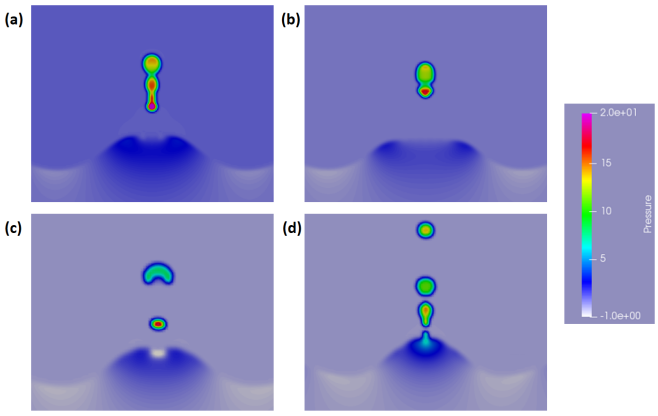


FIG. 6. Pressure Maps of different La's at specific time coordinates. (a) $La = 2 \times 10^3$(t=0.77) (b) $La = 2 \times 10^4$(t=0.80) (c) $La = 2 \times 10^4$(t=0.80) (d) $La = 2 \times 10^4$(t=0.80)

sion is an inward force acting on the droplets causing them to minimise surface energy by reducing their surface area. Curved surfaces are formed to minimise surface area. Therefore, spherical geometry allows for this to be achieved.

Bubble and droplet sizes can be determined both qualitatively and quantitatively. Quantitatively, the radius of each bubble and droplet displayed on figure 5 was measured. Due to their spherical nature, the volume of each bubble and droplet was computed using the volume of a sphere formula $V = 4\pi R_d^3/3$. R_d is the radius of the droplet. Taking $t = 0.80$ on figure 3 for instance, the radius of the first droplet of La 2×10^3 can be read off as 0.085 which gives a volume of 0.00257. This is in coherence with the experimental data obtained by Spiel [32] which shows that the radius of the first ejected drop is about one-tenth the initial radius of the collapsed bubble, R_o.

No correlation is exhibited between the droplet size and La which infers the possibility of viscosity having no effect on the sizes of droplets produced during the bubble bursting phenomena in liquid. More data of different La's is required to verify this conclusion. However, the trend of increasing size and velocity of bubble produced with La proves to be an important discovery as it can be linked with the number of droplets produced and their velocities.

Larger bubble size (represented by volume, V) means greater mass, m as $m = \rho V$. Both mass and velocity contributes to momentum. In coherence with Newton's Third Law of motion, an equal and opposite momentum is generated via the droplets. Therefore, as the downward momentum increases, the number and velocity of droplets increase to produce the equal and opposite momentum to counteract the momentum of the bubble.

Droplet velocity increases with La number. Gravitational effects on deceleration are negligible from the setting of the Bo number as discussed earlier. Based on Figure 7, a droplet position on the z coordinate can be traced with the moving peak with time. The diminishing peak within the velocity profile as time progresses proves that the droplet experiences deceleration upon being released by the jet. Due to the projectile of the droplet undergoing deceleration due to viscous effects as time progresses, the droplet velocity is taken as the initial velocity at the stage when the droplet is budding off the jet. For example, referring to t=0.50 of figure 5, the velocities of the second droplet of the La numbers in increasing order can be read off as 5.1, 6.8 and 9.9 respectively. Higher droplet velocity eventually leads to greater spatial heights

achieved by the droplet from the interface.

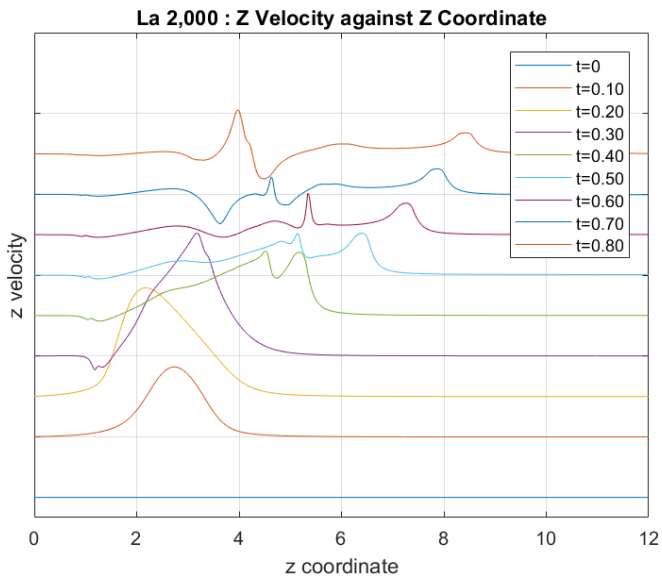


FIG. 7. Snapshots of z velocity profiles on an arbitrary scale along the z coordinate at La 2×10^3

IV. CONCLUSIONS

I have analysed the temporal evolution and dynamics of the three La numbers 2×10^3, 2×10^4 and 2×10^5 presented by the CFD numerical simulations. It was observed that lower viscosity gives a more robust droplet ejection dynamic, which promotes the release of compounds within the liquid to the surroundings. A higher number of droplets ejected and a greater travelled height from the interface reflects positively on the characteristics for mass transfer and evaporation such as the interfacial surface area and the exposure duration respectively. Higher droplet velocities also increase their rate of evaporation due to higher apparent velocity of air against the droplets. Based on the jet length analysis, the decrease in jet length from La 2×10^4 to 2×10^5 due to an earlier droplet ejection confirms a lower viscosity gives a higher rate of droplet ejection. Therefore, it would be recommended to use a liquid of lower viscosity to promote favoured mass transfer (e.g. champagne aerosols). On the contrary, a higher liquid viscosity would be preferred to control unfavoured mass transfer (e.g. polluted water, disease transmission). Similar dynamics may be observed when surface tension is varied (commonly with the use of surfactants) instead of viscosity provided the La number is the same. Further study has to be done to verify this.

V. ACKNOWLEGEMENTS

I would like to thank Ricardo Constante-Amores for the helpful discussions, technical support and training on the numerical simulation.

[1] Wu, Jin 1981 Evidence of sea spray produced by bursting bubbles *Science*, 324-326.

[2] Higgins F. Q. (1964) Bacterial aerosols from bursting bubbles. *PhD Thesis.*, Georgia Institute of Technology

[3] Slote L. 1976 Viral aerosols a potential occupationally related health threat in aerated wastewater treatment systems *J. of envi. health.*, **38** (5), 310-314.

[4] Tseng, Ruo-Shan and Viechnicki, John T and Skop, Richard A and Brown, John W 1992 Sea-to-air transfer of surface-active organic compounds by bursting bubbles *Journal of Geophysical Research: Oceans*, **97** (C4), 5201-5206.

[5] Poulain, S. and Villermaux, E. and Bourouiba, L. 2018 Ageing and burst of surface bubbles *Journal of Fluid Mechanics*, **851**, 636-671.

[6] Séon, Thomas and Liger-Belair, Gérard 2017 Effervescence in champagne and sparkling wines: From bubble bursting to droplet evaporation *The European Physical Journal Special Topics*, **226** (1), 117-156.

[7] Liger-Belair, Gérard and Cilindre, Clara and Gougeon, Régis D and Lucio, Marianna and Gebefügi, Istvan and Jeandet, Philippe and Schmitt-Kopplin, Philippe 2009 Unraveling different chemical fingerprints between a champagne wine and its aerosols *Proceedings of the National Academy of Sciences*, **106** (39), 16545-16549.

[8] Polidori, Guillaume and Jeandet, Philippe and Liger-Belair, Gérard 2009 Bubbles and Flow Patterns in Champagne: Is the fizz just for show, or does it add to the taste of sparkling wines? *American Scientist*, **97** (4), 294-301.

[9] Pozo-Bayón, Maria Ángeles and Santos, Mónica and Martín-Álvarez, Pedro J and Reineccius, Gary 2009 Influence of carbonation on aroma release from liquid systems using an artificial throat and a proton transfer reaction–mass spectrometric technique (PTR–MS) *Flavour and fragrance journal*, **24** (5), 226-233.

[10] Elisabeth Ghabache, Gérard Liger-Belair, Arnaud Antkowiak Thomas Séon 2016 Evaporation of droplets in a Champagne wine aerosol *Scientific Reports*, **6** 25148.

[11] Hewson, L and Hollowood, T and Chandra, S and Hort, J 2009 Gustatory, olfactory and trigeminal interactions in a model carbonated beverage *Chemosensory perception*, **2** (2), 94-107.

[12] Woodcock, AH and Kientzler, CF and Arons, AB and Blanchard, DC 1953 Giant condensation nuclei from bursting bubbles *Nature*, **172** (4390), 1144-1145.

[13] Beaumont, Fabien and Liger-Belair, Gérard and Polidori, Guillaume 2019 Three-dimensional model-

ing of complex swirling flows in champagne glasses: CFD and flow visualization *Acta Mechanica*, **230** (1), 213-224.

[14] Beaumont, Fabien and Popa, Catalin and Liger Belair, Gérard and Polidori, Guillaume 2014 Numerical modeling of bubble-induced flow patterns in champagne glasses *International Journal of Numerical Methods for Heat & Fluid Flow*, **24** (3), 563-578.

[15] Deike, Luc and Ghabache, Elisabeth and Liger-Belair, Gérard and Das, Arup K and Zaleski, Stéphane and Popinet, Stéphane and Séon, Thomas 2018 Dynamics of jets produced by bursting bubbles *Physical Review Fluids*, **3** (1), 013603.

[16] Lai, Ching-Yao and Eggers, Jens and Deike, Luc 2018 Bubble bursting: universal cavity and jet profiles *Physical review letters*, **121** (14), 144501.

[17] Lhuissier, Henri and Villermaux, Emmanuel 2012 Bursting bubble aerosols *Journal of Fluid Mechanics*, **696** , 5-44.

[18] Ghabache, Elisabeth and Antkowiak, Arnaud and Josserand, Christophe and Séon, Thomas 2014 On the physics of fizziness: How bubble bursting controls droplets ejection *Physics of Fluids*, **26** (12), 121701.

[19] MacIntyre, Ferren 1972 Flow patterns in breaking bubbles *Journal of Geophysical Research*, **77** (27), 5211-5228.

[20] Duchemin, Laurent and Popinet, Stéphane and Josserand, Christophe and Zaleski, Stéphane 2002 Jet formation in bubbles bursting at a free surface *Physics of fluids*, **14** (9), 3000-3008.

[21] Ghabache, Elisabeth and Séon, Thomas 2016 Size of the top jet drop produced by bubble bursting *Physical Review Fluids*, **119** (5), 051901.

[22] Seungwon Shin, Damir Juric 2005 Modeling Three-Dimensional Multiphase Flow Using a Level Contour Reconstruction Method for Front Tracking without Connectivity *Journal of Computational Physics*, **180** (2), 427-470.

[23] Seungwon Shin, S. I. Abdel-Khalik, Virginie Daru, Damir Juric 2005 Accurate representation of surface tension using the level contour reconstruction method *Journal of Computational Physics*, **203** (2), 493-516.

[24] Seungwon Shin, Damir Juric 2007 High order level contour reconstruction method *Journal of Mechanical Science and Technology*, **21** (2), 311–326.

[25] Seungwon Shin, Damir Juric 2009 Simulation of droplet impact on a solid surface using the level contour reconstruction method *Journal of Mechanical Science and Technology*, **23** (9), 2434–2443.

[26] Seungwon Shin, Jalel Chergui, Damir Juric 2017 A solver for massively parallel direct numerical simulation of three-dimensional multiphase flows *Journal of Mechanical Science and Technology*, **31** (4), 1739–1751.

[27] Krishnan, S., Hopfinger, E.J. and Puthenveettil, B.A. 2017 On the scaling of jetting from bubble collapse at a liquid surface *Journal of Fluid Mechanics*, **822** 791-812.

[28] Arthur Mason Worthington and R. S. Cole 1897 V. Impact with a liquid surface, studied by the aid of instantaneous photography *Philosophical Transactions of the Royal Society A*, **189** 137–148.

[29] J. M. Gordillo, J. Rodríguez-Rodríguez 2019 Capillary waves control the ejection of bubble bursting jets *Journal of Fluid Mechanics*, **867**, 556-571.

[30] Gañán-Calvo, Alfonso M. 2017 Revision of Bubble Bursting: Universal Scaling Laws of Top Jet Drop Size and Speed *Phys. Rev. Lett.*, **119** (5), 204502.

[31] U-Ser Jeng, Levon Esibov, Lowell Crow and Albert Steyerl 1998 Viscosity effect on capillary waves at liquid interfaces *Journal of Physics: Condensed Matter*, **10** (23), 4955.

[32] Donald E. Spiel 1995 On the births of jet drops from bubbles bursting on water surfaces *JOURNAL OF GEOPHYSICAL RESEARCH*, **100** (C3), 4995-5006.

Towards a More Sustainable Lithium-Ion Battery Future: Recycling LIBs from Electric Vehicles

Aniket Chitre[1] & Daniel Freake[1]

There is going to be a huge increase in spent lithium-ion battery (LIB) waste over the next two decades, hence, to minimise supply chain risks and environmental & safety concerns, LIBs need to be effectively recycled. The main challenge to recycle LIBs arises from their complex chemistry and variability, both in terms of cathode chemistries and physical shapes & sizes. The current global recycling rate of LIBs is estimated to be around 50%, but processes need to be vastly improved if we are to conserve rare resources and handle much larger throughputs of spent LIBs. Present processes use pyrometallurgical and/or hydrometallurgical recycling methods, however, pyrometallurgy is largely to be viewed as a process of the past due to its lower material recovery rate. This makes it less economically viable for recycling newer battery chemistries with lower cobalt contents. As such, this paper focuses on hydrometallurgy, including improvements which can be made in sorting batteries and using alternative water-soluble binders to improve battery recyclability. This paper promotes the adoption of a holistic design approach for LIBs that includes ease of end-of-life recyclability.

Introduction

Electric vehicles (EVs), powered by lithium-ion battery technology, are widely anticipated as being a leading part of the solution to decarbonise the transportation sector (Hill, et al., 2019). Currently there are just over 5 million EVs on the road and the International Energy Agency (IEA) projects that the global EV stock will exceed 130 million by 2030 (IEA, 2019). Consequently, this will lead to significant LIB waste management issues. One study predicted that the cumulative LIB waste from EVs will build up to anywhere in between 0.33 and 4 million tons between 2015 and 2040 (Richa, et al., 2014). Currently, LIB recycling is underdeveloped on an industrial scale. This paper presents a review into various recycling technologies and highlights proposals that address key process challenges for recycling LIBs.

Recycling reduces the extraction of virgin ores, thus extending the life of these resources. The metals found in spent LIBs can generally be found at high concentration levels, sometimes higher than those found in the processing of concentrates of natural ores themselves (Zeng, et al., 2014). Lithium, cobalt and manganese, all used in LIBs, have been cited by the U.S. Government as critical metals (US Department of the Interior, 2018), meaning they are of high economic importance and in scarce supply. In 2017 30% of cobalt demand came from battery applications alone and this is predicted to rise to 53% by 2025 (McKinsey & Company, 2018). Cobalt is a key element in the cathodic material of a battery, however, estimates are that only 50 to 100 years remain before all known cobalt reserves are depleted (Hunt, et al., 2015). Such supply chain risks pose a serious threat to the LIB industry; therefore, recycling is essential for sustainable resource management.

The supply chain risks are exacerbated as lithium and cobalt reserves are geographically concentrated in select areas of the world, some of which are politically volatile. For instance, the Democratic Republic of Congo (DRC) is home to more than 60% of the world's mined cobalt output (Larsen, 2019). Another example is China's recent efforts to control foreign lithium and cobalt mining supplies (South China Morning Post, 2018). Such oligopolistic control would pose risk to the industry. Furthermore, many of the mining practices in the DRC are highly unethical, with one third of the country's cobalt supply produced through the use of child labour in artisanal mines (Financial Times, 2019). Effective recycling of LIBs would enable countries to secure a domestic supply of the valuable materials needed to sustain the industry.

Environmental and safety concerns of LIBs also demand that effective recycling strategies be developed. The initial extraction of the metals is highly energy intensive, so using recycled materials can reduce energy consumption and emissions (Gaines, et al., 2018). To obtain 1 ton of lithium requires 250 tons of mineral or 750 tons of brine, compared to 28 tons of used batteries (Larcher & Tarascon, 2014). Additionally, LIBs can be a significant fire-hazard and can release toxic fumes from the electrolyte if ill-disposed of (Harper, et al., 2019). An effective recycling strategy must also mitigate these safety risks.

Clearly, there is a strong case for the need to develop wide-scale LIB recycling. In the following section, the paper provides a background on the lithium-ion cell, outlines the challenges faced for wider-scale adoption of LIB recycling and reviews current recycling technologies.

[1] Department of Chemical Engineering, Imperial College London, U.K.

Background

Structure and function of lithium-ion cells

A LIB has 4 basic components: cathode, anode, electrolyte, and separator. A reversible chemical reaction occurs between the electrodes of a rechargeable cell and this has an ionic and electronic component. The electrolyte transports the ionic component inside a cell and forces the electronic component to traverse an external circuit (Goodenough, 2018).

The electrolyte in a LIB consists of a lithium salt (usually $LiPF_6$) dissolved in an aprotic organic solvent. The separator is a component that prevents the contact between the two electrodes and facilitates the flow of electrons through the electrolyte – it is a thin, porous plastic film (usually polyethylene or polypropylene) (Gaines, et al., 2018).

The physics of the cell works by shuttling lithium ions between higher and lower potential energy states of the anode and cathode respectively (Winslow, et al., 2018). The anode is typically made up of a copper current collector sheet that is coated with a graphite layer. The anode is usually similar even in different types of LIBs, however, the cathode, a lithium oxide bound on an aluminium plate, can vary considerably. There are various possible forms for the lithium oxide, such as lithium cobalt oxide, $LiCoO_2$ (LCO), which was the first-generation cathodic material. However, a wide variety of cathodic chemistries are now present, and this poses a significant challenge for recycling. Some example chemistries include: NMC ($LiNiCoMnO_2$), NCA ($LiNiCoAlO_2$), LMO ($LiMn_2O_4$), and LFP ($LiFePO_4$). Far more variability is present with the cathodic material as it is the key component where research is focused to improve performance and drive down costs (Gaines, et al., 2018). Figure 1 shows a summary of how different battery chemistries compare with each other.

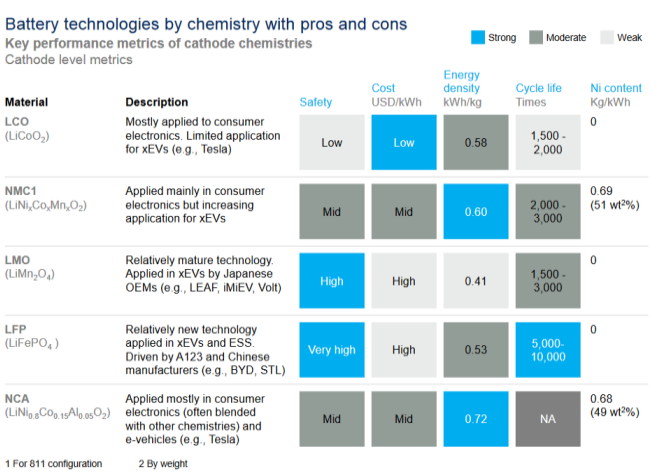


Figure 1: Comparison and description of cathode chemistries in terms of safety, cost, energy density, cycle life and nickel content. (McKinsey & Company, 2018)

An important part of the electrodes is the binder, which is used to fix the active electrode materials to the metal current collector (aluminium and copper plates). Due to the reactive environment inside a battery the binder must withstand both heat and electricity. Polyvinylidene fluoride (PVDF) is the binder material of choice for its thermoresistant and nonreactive properties (Zeng, et al., 2014).

LIB chemistries used in EVs

Figure 1 shows that the main four metrics along which LIBs are compared are safety, cost, energy density, and cycle life. There is no single standout cathode composition that dominates global markets – it largely depends on application (Arambarri, et al., 2019). Here, the focus is on the EV market where NMC and NCA are the two most common chemistries, yet there is still variability. For example, among NMCs themselves there are several different compositions. Researchers have been continuously trying to improve battery performance and reduce cost – the latter primarily through substituting out very expensive cobalt from the cathode. NMC cathodes with different Ni-Mn-Co compositions have been around for nearly 20 years now, with several key publications and patents published in the mid-2000s (Myung, et al., 2017). The first developed NMC cathode was NMC 333 ($LiNi_{1/3}Co_{1/3}Mn_{1/3}O_2$) and over time manufacturers have attempted to substitute the cobalt with more nickel. This has led to the formation of NMC 532 and 622 chemistries. Earlier this year the Chinese battery manufacturer, CATL, was the first to commercialize NMC 811 (80% nickel, 10% cobalt, 10% manganese) batteries (Deign, 2019).

A 2010 report from consulting giant BCG identified NMC, NCA and LFP as the most promising candidates for EV applications. With hindsight we can verify in the past decade that the majority of EVs use some sort of NMC chemistry (the Nissan Leaf, Chevy Volt, BMW i3 etc.), the notable exception which uses NCA cells being Tesla (Oberhaus, 2019). LFPs, which are intrinsically less prone to thermal runaway, were also expected to be part of the mix. However, this was not the case due to their lower energy density, as manufacturers seem to have prioritised driving range – a directly tangible consumer criterion. The key takeaway from this subsection is that LIBs are within the fluid phase of their product life cycle and this technology is rapidly evolving.

Current challenges in recycling LIBs

The current recycling rate of LIBs is estimated to be around 50% (Gupta, 2019), compared to lead-acid batteries, the most widely recycled consumer product, which have a 99% recycling rate in the United States

(Gaines, 2014). A comparison with lead-acid batteries highlights some of the key challenges to recycle LIBs. Lead-acid batteries are widely standardized – essentially all the manufacturers use the same raw materials: lead, lead oxide, and sulfuric acid in a polypropylene case, whereas there is a wide variety for LIBs.

Firstly, there is significant variability with complex cathode chemistries. This creates challenges on both the technological and economic fronts. Depending on the battery chemistry, different processes and chemicals need to be used. Studies have pointed out that a more flexible and universal process is required for the recovery of different types of LIBs (Yao, et al., 2018). Furthermore, economic analysis has shown that recycling is typically not financially feasible unless there are significant cobalt and nickel contents to be recovered in the waste-stream (with pyrometallurgical processes), whereas lead-acid batteries are always profitable to recycle (Gaines, 2014). One study concluded that a waste stream of LMO batteries ($LiMn_2O_4$) would only be worth \$860/ton to recycle, whereas, LCO batteries ($LiCoO_2$) would be worth \$8900/ton (Wang, et al., 2014). Therefore, the feasibility of a LIB recycling process is currently highly dependent on the cathode chemistry.

The components of the battery must be packed in an outer casing. A typical cell is constructed by winding, stacking or folding together strips of cathode, separator, and anode and packing them tightly into casing (Gaines, et al., 2018). The fact that the cells are packed into different cell types, typically prismatic, planar or cylindrical, adds to the complexity of recycling them (Harper, et al., 2019). The cells are then arranged into modules and built up to battery packs used in EVs. There is no standardisation as there is for lead-acid batteries.

Processes do exist for LIB recycling and these are reviewed next, before moving on to suggest improvements to facilitate better recycling rates.

Pyrometallurgical process

Pyrometallurgy uses high temperatures to break down spent LIBs. In this process LIB cells are sent to a shaft furnace. The process is split up into three separate steps or temperatures (Huang, et al., 2018). The first step is pre-heating, where the batteries are heated up at a relatively low temperature of around 300 °C in order to evaporate the electrolyte. This reduces the possibility of hazards in the following steps. The second step is plastics pyrolysis, which utilises higher temperatures of around 700 °C to burn off the plastics and to lower the energy consumption of the final step. The final step is the smelting and reducing step where all the materials are smelted into either an alloy or a slag. Depending on the type of LIB being recycled, the alloy can contain copper, cobalt, nickel, and iron, while the slag can contain lithium, aluminium, calcium, and iron. The recovered alloy is then sent for further treatment to recover the valuable metals while the slag is usually disposed of or used as a building material (Xiao, et al., 2017). This process lets off dangerous gases, most notably hydrogen fluoride when the electrolyte is evaporated. These gases need to be captured and cleaned to minimise environmental damages. The alloy usually requires several different treatment steps in order to extract the transition metal from it. These steps can be hydrometallurgical or bio-metallurgical (Yun, et al., 2018).

Hydrometallurgical process

Hydrometallurgical LIB recycling processes can be split into three different steps; pre-treatment, leaching, and metal recovery (Yao, et al., 2018). The batteries first need to be opened and pre-treated to isolate the electrode materials, from which the key materials are recovered. During pre-treatment the LIBs need to first be discharged, to avoid fires and short-circuiting. This can be done using a discharger or immersing the batteries in an ionic solution. Next the plastic and metal coatings need to be stripped off the battery. This is currently done manually with hand tools or on a larger scale the batteries are crushed and sieved. Neither the manual nor machinal approach are very efficient. Further discussion is provided under the 'labelling and packaging' section below about the future use of advanced robotics for industrial-scale LIB disassembly. With the batteries opened, the cathodic material needs to be separated from the aluminium plate it is bound to (using PVDF binder). This is most commonly done through dissolution of the PVDF binder using a solvent such as *N*-methyl pyrrolidone (NMP) (Harper, et al., 2019). This issue is elaborated in much greater depth in the 'binder material' section below. Then the leaching step takes the cathodic material, now separated from the foil and in powder form, and forces the metal ions out into solution. The leaching can be carried out by organic or inorganic acids, bases, or bacteria. The final step is to recover the metals from the leached solution. There are many methods of doing this and in general more than one needs to be used in order to maximise recovery. Examples of the recovery methods are chemical precipitation, electrochemical deposition, and solvent extraction. Chemical precipitation uses precipitants containing anions that can form precipitates with the metal ions in solution. This method works best for simple solutions but can be used for more intricate ones by using several precipitants in parallel. Electrochemical deposition separates metals from solution through the difference in their electrode potential. Solvent extraction is a widely implemented

method which uses solvents to selectively extract metal ions (Yao, et al., 2018). The main discussion elaborates on appropriate processes and reagents for recycling the LIB chemistries found (and to be found) in EVs.

Direct recycling

For the sake of completeness, direct recycling is included, despite this technology not being ready for commercial deployment yet. Direct cathode recycling recovers and reconditions the powdered cathode material for use in subsequent batteries (Ciez & Whitacre, 2019). The motivation behind this process is illustrated through the cost analysis of the 2019 Nissan Leaf battery. This battery costs ~ $6500-8500 new. The cathodic materials of this battery in their pure form are worth ~$400, while the cost of the cathode material is ~$4000 (Harper, et al., 2019). The advantage of 'upcycling' the electrode in this manner is that a lot more of the electrode value is retained. Direct recycling shares similar initial steps with hydrometallurgical processes, but subsequently chemical treatment steps are not utilised, as the goal is to maintain the integrity of the cathode. Instead the cathode is physically recovered and relithiated for reuse. However, the process specifics are largely undeveloped and it is not clear whether the recovered materials will match the long-term performance of fresh ones (Lv, et al., 2018).

Comparison of processes

All currently existing industrial LIB recycling facilities employ either pyro- or hydrometallurgical processes, or a combination of the two (Lv, et al., 2018). The focus of this paper is to review industrially scalable recycling technologies and suggest routes and advancements for these processes, hence technologies which are still at the lab-scale are not considered further. Pyrometallurgical processes were the first generation of industrial LIB recycling facilities, offering the advantage that LIBs of different varieties could be recycled together without the need for any pre-sorting. However, especially with EV and battery manufacturers attempting to move away from cobalt chemistries, pyrometallurgical processes are finding it increasingly difficult to remain profitable. Additionally, these processes are very energy intensive, produce significant quantities of hazardous gases, and are not able to effectively recover a lot of the materials of LIBs (Huang, et al., 2018). Several studies have concluded that hydrometallurgical methods have obvious advantages over pyrometallurgical ones. They are more environmentally friendly, use milder reaction conditions, and have higher recovery efficiencies for valuable metals (Lv, et al., 2018). Hydrometallurgical processes have their own drawbacks, such as requiring very expensive solvents and several complicated processing steps, yet they are currently the most suitable method for recycling spent LIBs (Yao, et al., 2018). A report by the Swedish Energy Agency provides two pieces of evidence supporting this claim. Firstly, an overwhelming majority of research effort is focused into hydrometallurgical processes over any other technology. Secondly, the recycling processes in China and Korea (where LIB recycling is most prominent) use "almost exclusively hydrometallurgical processes" (Melin, 2019). The rest of the paper considers hydrometallurgy for LIB recycling. For any processes other than pyrometallurgy the LIBs need to be physically opened. Therefore, the main discussion starts with a focus on 'labelling and packaging' of LIBs and then considers the binder material of choice, which currently presents a key process bottleneck. Finally, downstream separation issues are considered.

Main Discussion

This study focuses on the technological aspect of LIB recycling. To have high recycling rates across society, rules and regulations will need to be put in place for proper LIB disposal, collection, and transfer to recycling facilities. For example, the existence of such legislation for lead-acid batteries in the United States is one of the key drivers behind their high recycling rate (Gaines, 2014). However, this is considered outside the scope of this paper. It is assumed a priori that spent LIBs have been successfully collected and the focus is on the technical aspects of the problem studied.

Labelling and packaging

Batteries need to be sorted for two reasons: i) batteries of physically different shapes and sizes need to be opened differently; and ii) batteries of different chemistries need to be treated differently. Standardised labelling would be necessary for effective sorting of batteries. Several groups, such as the Battery Recycling Committee of the Society of Automotive Engineers (SAE) or the Battery Association of Japan, have developed labels which describe the battery's size, chemistry and origin (Gaines, et al., 2018). Figure 2 is a modified prototype of such a label.

In order to effectively separate a mixed battery stream and then take them apart, advances need to be made in AI and robotics technology, as the cost to manually disassemble LIBs is too great to be economically feasible. Robots using computer vision (AI) would have the ability to perform the required separation as demonstrated by the various robots in place using similar technology to separate plastics (Kite-Powell, 2017). The labelling shown below would aid the robots in their task, but diagnosis of each individual LIB would have to conducted to evaluate cell health.

Currently, LIBs are assembled primarily with manufacturing costs and safety in mind, and not disassembly. This means that a lot of battery packs use strong adhesive substances or welds to ensure that the cells do not move, instead of using manual attachments such as bolts and screws. This makes LIB packs difficult to take apart as high temperatures or heavy-duty mechanical equipment is required in order to weaken the adhesive or to break the weld. Policies which force LIB manufacturers to facilitate disassembly need to be put in place. This could entail only allowing use of manual attachments such as nuts and bolts.

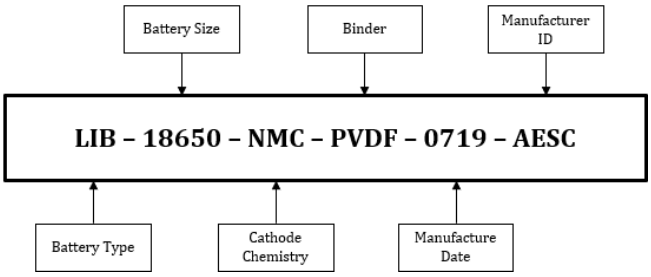


Figure 2: Battery label prototype, developed by Daniel Freake

Once the battery has been disassembled, what is left is the cathodic and anodic materials bound to the aluminium and copper plates respectively. How easy it is to separate the electrode materials from their current collector plates depends on the binder material used and this is discussed next.

Binder material

Other than with pyrometallurgy, other LIB recycling routes require the electrode materials to be separated from the current collector plates which they are bound to. Separating the active cathodic material is of most interest as it contains most of the high value metals which are worth recovering. As mentioned above there are several downsides with the pyrometallurgical processes, hence as society moves towards other processes the role of the binder material and the ease of detaching the electrode materials from the metal foils will become more important.

As mentioned in the background section polyvinylidene fluoride (PVDF) is the most commonly used binder material in commercially available LIBs (Zeng, et al., 2014). Initially little attention was directed towards alternative binder materials as PVDF had been performing well and the scale of LIB production was rather low, hence some of the issues which will be raised below were of secondary importance (Bresser, et al., 2018). However, with issues of process scalability now being important for the LIB industry, this is becoming a key research area.

The choice of binder material impacts two crucial factors: the ease of processability (both in the manufacturing and EOL recycling stages) and the cell performance (Shi, et al., 2017). Within the frame of this study, most of the focus is on the processability.

When PVDF is used as the binder, the most common pre-treatment route for liberating the graphite and metal oxides from the Cu and Al current collectors is the use of sonication in a solvent such as *N*-methyl pyrrolidone (NMP). Then thermal treatment is used to decompose the binder, leaving only the electrode materials behind for material recovery (Harper, et al., 2019).

There are several problems with this process. First of all, it is relatively slow – it takes around 3 hours, although with the assistance of ultrasonication (use of higher energy sound waves) this process time can be brought down to about 1.5 hours (Harper, et al., 2019). However, this is still too slow for the development of a continuous-flow process. There is also the issue that the required solvent-to-solid mass ratio of 10:1 will not be viable on a commercial scale (Harper, et al., 2019), since this is an expensive solvent. Furthermore, both NMP and PVDF are hazardous - NMP is teratogenic and an irritant, while PVDF is mutagenic and teratogenic. Industry needs to look towards using more benign chemicals. In response to awareness of these hazards, the EU has introduced new legislature, as an annexe to REACH (Regulation 907/2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals), which states that after May 9th, 2020 *"NMP shall not be used as a substance on its own or in mixtures in a concentration equal to or greater than 0.3% unless manufacturers and downstream users take the appropriate risk management measures [to keep exposure below certain specified limits]"* (European Chemicals Agency, 2019).

The alternative to using NMP to dissolve a PVDF binder is to use a water-soluble binder. This transition would be coined as moving from organic-based to aqueous-based electrode processing in literature. In addition to mitigating the above outlined issues, aqueous-based electrode processing presents the following advantages - it leads to a substantial decrease in process CO_2 emissions and is much more cost-effective. NMP costs \$1-3/kg compared to roughly \$0.015/L for water. PVDF costs \$8-10/kg compared to carboxymethyl cellulose (CMC), the most commonly used aqueous binder, costing \$2-5/kg (Bresser, et al., 2018).

When identifying alternative binders, their environmental impact can be assessed according to three criteria: processability, chemical composition, and natural availability. Processability in this context means the binder can be processed in environmentally benign solvents such as water or ethanol. An environmentally

friendly binder needs to be fluorine free, as the decomposition products of fluoropolymers are extremely toxic and known to affect the ozone layer (Jiang & Pickering, 2016). And finally, natural availability refers to whether the material is synthetic or bio-derived – the latter being the desired, more sustainable option. Figure 3 shows various classes of binders, as considered in a thorough review on the matter (Bresser, et al., 2018).

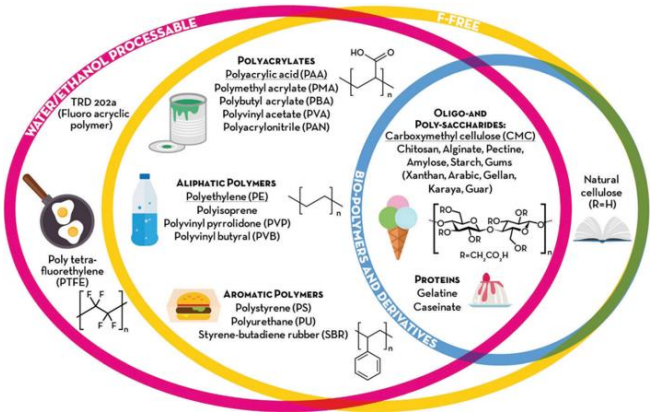


Figure 3: Overview of different classes of binders as categorised according to their processability (water and/or ethanol processable), chemical composition (F-free) and natural availability (bio-polymers and derivatives) (Bresser, et al., 2018)

The review by Bresser et al. shows that carboxymethyl cellulose (CMC) is the most promising water-soluble binder (Bresser, et al., 2018). Despite all the evidence pointing towards the adoption of water-soluble and environmentally friendly binders, CMC has only been incorporated into anodes due to distinct technical challenges for aqueous-based cathode processing. The main issue impeding the implementation of water-soluble binders to cathodes is the reactivity of most commercially employed cathode materials with water. When lithium transition metal oxides are dispersed in water the lithium leaches out of the active material host structure. This leads to the concurrent formation of $LiOH_{(aq)}$ species, causing an increase in pH, which has a further knock-on effect of corrosion of the aluminium current collector plate (Bresser, et al., 2018). This can be reasoned through an inspection of the Pourbaix diagram for Al – the increase in pH pushes the metal out of its immune/passivation regimes and into its corrosion regime. The lithium leaching reduces the amount of electrochemically active lithium available in the cell, thus hampering electrochemical performance.

Guerfi et al. began the first investigations of incorporating water-soluble binders in LIBs by using CMC for the synthesis of LFP cathodes in 2007 (Guerfi, et al., 2007). Initial work started with LFP cathode chemistries because they are less sensitive to water and moisture. The strong Fe-P-O bonding makes this cathode material more inert (Bresser, et al., 2018), and hence LFPs mitigate the delithiation and corrosion issues that stem from aqueous-based cathode processing. As discussed in the background section, LFPs have not seen major adoption in the EV industry due to their lower energy density, especially compared to the preferable NMC and NCA chemistries - LFPs have a 12% lower energy density than NMCs (Figure 1). Policymakers must consider whether a certain performance drawback is acceptable as a trade-off for more responsible environmental stewardship.

Recent breakthroughs in the past couple of years have been achieved by incorporating water-soluble binders into NMC cathodes. Two bodies leading research on this front are the Oak Ridge National Laboratory (ORNL) in the United States (Wood, et al., 2018) and the European Union's GREENLION research project (IK4-CIDETEC, 2016). To address the corrosion of the aluminium current collector three main solutions are being investigated: i) a corrosion-resistant carbon coating of the aluminium foil; ii) using a co-binder (e.g. polyurethane) which inhibits corrosion; iii) adding acid to adjust (i.e. lower) the slurry pH to keep the aluminium in its passivation regime (Bresser, et al., 2018). As this is an active research area, these routes are not fully understood yet. Additionally, to take these technologies from the laboratory to industry research needs to be focused on the scale-up. There are also key process engineering challenges present, particularly with optimising the mixing regime for the electrode slurry and the subsequent drying process – these have direct implications for electrode performance.

Loeffler et al. with funding from the European GREENLION project managed to manufacture two cells using aqueous-based electrode processing. One used a sodium-carboxymethylcellulose (Na-CMC) binder with styrene-butadiene rubber (SBR) (a commonly used 'thickener' in electrode manufacturing processes) for an NMC 333 cathode. The other used Na-CMC with a water-borne latex binder (Loeffler, et al., 2017). Another investigation from ORNL used a similar water-soluble CMC and latex binder combination to prepare an NMC 532 cathode. Results from both these experiments showed comparable or better electrochemical performance than equivalent cathode types formulated with PVDF binder. Cells from the European study showed a remarkable capacity retention – only fading 0.004% per cycle (Loeffler, et al., 2017).

Further work is required in this area, however, it looks like a promising transition could soon be made to water-soluble binders. This would bring substantial economic and environmental benefits (Wood, et al., 2018).

Downstream hydrometallurgical separations

This section considers the specific hydrometallurgical processes required to recycle LIBs from EVs. As explained in the background section, NMC and NCA are the two most prevalent LIB chemistries in EVs. Additionally, recycling of LFPs has been considered because these are expected to be adopted into the market, especially in China, due to the continued increase in cobalt prices (Sofiane & Ravichandran, 2019). Figure 4 depicts possible hydrometallurgical recycling routes for the three battery types, and then an estimation of the cost, environmental impact and process time is summarised in Figure 5 to rank the illustrated processes from best to worst.

An abundance of literature is focused on hydrometallurgical recycling processes (Lv, et al., 2018), (Yao, et al., 2018), (Huang, et al., 2018). However, most of the work is focused on the LCO cathode chemistry, as this was used in the first generation of LIBs to be commercialised. LIBs with LCO cathodes are not used for EVs as they contain too high a cobalt content and are unsafe for application in EVs (as shown in Figure 1). NCA batteries have only been recently commercialised and LFPs have not been that widely adopted (except in buses), so there are fewer studies focusing on their recycling (Arambarri, et al., 2019). The studies that have been performed lack some of the sensitivity analysis that the processes of more established battery types have. For this reason the depicted processes for LFP (Zheng, et al., 2016) and NCA (Joulié, et al., 2014) batteries are not corroborated by other papers but a possible path for their recycling is demonstrated. In contrast NMC batteries have a variety of reported recovery processes because they are older and more widely used. NMC in its 333 form is the oldest NMC battery and hence most of the studies focus on this type.

There were a variety of methods to choose from for the recycling of NMC LIBs, but a relatively straightforward pathway that prioritised recovery was chosen. Many acids have been researched for the leaching of NMC cathodes, such as citric acid, acetic acid and sulphuric acid. Sulphuric acid was chosen in this case because it offers a good balance between leaching time and solid-liquid mass ratio. The recovery of the metals from the leached solution can again be performed by a variety of chemicals. The specific pathway chosen was taken from Zheng, et al. and Wang, et al.

Figure 4 is structured in four vertical steps. The hydrometallurgical processes all share a common route, independent of the cathode type, to open the battery and then pre-treat it with NMP to isolate the cathodic material away from the aluminium foil. The point of interest is therefore considered from the leaching process onwards. This is the first step shown on the schematic diagram below – the illustrated processes use inorganic acid leaching. The next two steps represent two core chemical treatment steps for metal recovery. Several steps of the hydrometallurgical process have been streamlined to show the core reagents used. Finally, the bottom of the figure shows a final recovery step (for NMC 333 only) and the recovered products in their final form.

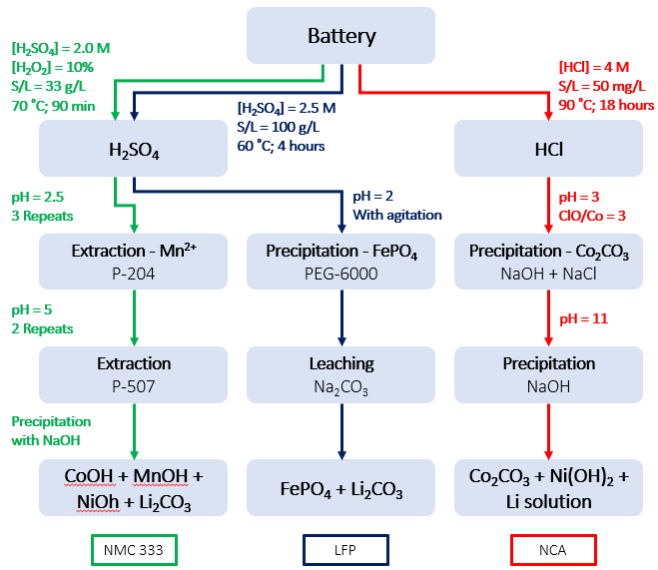


Figure 4: Downstream hydrometallurgical processes for 3 LIBs (NMC 333: (Zheng, et al., 2016) *and* (Wang, et al., 2018)*, LFP:* (Zheng, et al., 2016)*, and NCA:* (Joulié, et al., 2014)*), developed by Daniel Freake. Note: the process begins from powdered cathodic material.*

	NMC 333	LFP	NCA
Cost	• • •	•	• •
Environment	•	• • •	• •
Time	•	• •	• • •

Figure 5: Comparison of cost, environmental impact and time of recycling for 3 LIBs (NMC 333, LFP, and NCA), developed by Daniel Freake. 1 dot: best, 2 dots: intermediate, and 3 dots: worst.

Price-data was found from Sigma Aldrich to conduct a cost-analysis of the three LIB processing routes. CapEx is not factored into the cost-estimates in this paper, the "cost rating" in Figure 5 is based primarily on the cost of chemicals needed during the recycling process. Since P-204 (dicotyl phosphate) and P-507 (phosphoric acid) are by far the most expensive chemicals used, the recycling cost of NMC was marked as being 'the worst'. The recycling of LFP requires only relatively cheap chemicals with the most expensive being sodium carbonate and

hence was marked as being 'the best'. The environmental impact was assessed through the CO_2 emissions mitigated by the recycling processes (Ciez & Whitacre, 2019). The time taken to recycle the specific battery was based mostly on the leaching time because it was the part of the process with the most widely reported time. The point to be demonstrated here is that the ease of recyclability and environmental friendliness of such a process is largely dependent on the battery type being recycled. This needs to be holistically considered as society tries to move towards LIBs and EVs in order "save the planet". Proper process modelling and more detailed calculations are required to gauge the impact of the different hydrometallurgical processes depicted dependent on cathode type.

Figure 4 shows that hydrometallurgical processing routes are very specific. If one considers that the average lifetime of a process plant (i.e. recycling facility) is at least 20 years, versus the average lifetime of a LIB in an EV of 8-10 years (Boston Consulting Group, 2010), and then factors in the fast-moving nature of LIB technology, it is likely that a recycling facility's LIB feedstock will vary significantly in composition over the course of its lifetime. Hence recycling facilities will need to be flexible. Plant design should be constructed modularly, so, if necessary, processes can be modified to deal with different cathode chemistries. For example, Tesla is using NCA cells today, but is expected to transition to NMC cells after a breakthrough with a partnered research institution (Dalhousie University), which managed to develop a LIB which can "last a million miles" (over double the current capacity) (Oberhaus, 2019). Similarly, BMW currently uses NMC 333 in its i3, but is planning to shift to NMC 622 in 2021 and further transition to NMC 811 by 2025 (Behrmann, et al., 2018). The implication here is that process flexibility is of paramount importance to recycling facilities.

Conclusion

EVs are being adopted as a solution to one sustainability issue, however, if we do not develop more effective LIB recycling strategies then we shall drive ourselves towards another one.

Current LIB recycling facilities use pyrometallurgical or hydrometallurgical processes, or a combination of the two. However, pyrometallurgical processes will become a process of the past. One of the key complications of recycling LIBs is the large variety of them present. In terms of cathode chemistry alone, there are several main chemistries (LCO, NMC, NCA, LFP etc.), and within these there are several sub-chemistries (e.g. NMC 333/532/622/811). Battery and EV manufacturers are looking towards decreasing the cobalt content in batteries, as its price is high and volatile. As society moves away from the first generation of LIBs (LCO chemistries), pyrometallurgical processes will become increasingly infeasible from an economic perspective because they do not provide efficient material recovery rates. Furthermore, pyrometallurgical processes are highly energy intensive and produce a lot of hazardous fumes which need careful treatment. Hydrometallurgical processes, despite their own drawbacks, are widely seen as the way forward. This is primarily due to their potential to recover products more selectively.

Industry and governments can learn how to successfully increase LIB recycling rates from case studies of waste management success stories. For example, lead-acid battery recycling is very commonplace (with a 99% recycling rate in the US). A major factor behind this is the standardisation of lead-acid batteries. Unfortunately, the LIB industry is not ready for such standardisation because they are still within the fluid phase of their product life cycle. Standardisation would restrict the technological progress which is necessary to bring EVs to parity with internal combustion engine vehicles.

Even though policymakers should not attempt to standardise LIB chemistries, it is advisable for them to regulate how the batteries are assembled. The first step of a hydrometallurgical process is to disassemble and strip open the battery and batteries should be designed to make them easier to disassemble, while of course ensuring they are still safe. For wide-scale LIB recycling, manually stripping the batteries will not be a viable option – this will have to be done by machines. Currently strong adhesives and/or welds are used to join battery components. However, this makes disassembly challenging, as high temperatures or heavy-duty mechanical equipment is required. Instead, this paper urges manufacturers to look towards using nuts & bolts and other such components which can be more easily handled by advanced robots.

A case study which can be drawn upon is plastic recycling. This sector demonstrates similar variability to LIBs. Robotics and computer vision AI algorithms are used for sorting and separation. Similarly, efforts should be directed towards advancement of AI and robotics for LIB recycling. To aid the use of AI for sorting a standardised labelling system for LIBs is suggested.

Next in a hydrometallurgical process the batteries are pre-treated before the electrode materials of interest can be isolated. Despite the binder material seeming innocuous, PVDF requires the use of large quantities of hazardous NMP solution to detach the cathode powder from the aluminium current collector foil. This is a key process bottleneck, but recent breakthroughs in the use

of water-soluble binders are promising. Further research needs to be done on aqueous-based cathode processing, especially on the scale-up aspect. However, the use of water-soluble binders is the direction the LIB industry needs to head in, which would bring a range of economic and environmental benefits to the manufacturing process.

Once the electrode materials have been isolated further chemical treatment steps need to be carried out to recover the desired materials downstream. These steps are very specific and dependent on the LIB type being recycled. Since LIB technology is evolving on a faster timescale than the current lifetime of battery recycling facilities, it's important that the chemical plants are designed in a modular fashion with process flexibility in mind. The plants may need to be retrofitted as hydrometallurgical processes evolve to recycle LIBs with different chemical compositions.

Overall, work is needed on both the policy and technological side for a more sustainable mobility powered by LIBs. In particular, manufacturers need to take EOL recyclability into consideration when they design LIBs.

References

1. Arambarri, J. et al., 2019. Lithium ion car batteries: Present analysis and future predictions. *Environmental Engineering Research,* 24(4), pp. 699-710.
2. Behrmann, E., Farchy, J. & Dodge, S., 2018. *Hype Meets Reality as Electric Car Dreams Run into Metal Crunch.* [Online] Available at: https://www.bloomberg.com/graphics/2018-cobalt-batteries/ [Accessed 26 November 2019].
3. Boston Consulting Group, 2010. *Batteries for Electric Cars,* s.l.: s.n.
4. Bresser, D. et al., 2018. Alternative binders for sustainable electrochemical energy storage - the transition to aqueous electrode processing and bio-derived polymers. *Energy & Environmental Science,* Volume 11, pp. 3096 - 3127.
5. Ciez, R. E. & Whitacre, J., 2019. Examining different recycling processes for lithium-ion batteries. *Nature Sustainability,* Volume 2, pp. 148-156.
6. Deign, J., 2019. *How the Battery Sector is Looking to Improve Lithium-Ion.* [Online] Available at: https://www.greentechmedia.com/articles/read/how-the-battery-sector-is-looking-to-improve-lithium-ion [Accessed 25 November 2019].
7. Dewulf, J. et al., 2010. Recycling rechargeable lithium ion batteries: Critical analysis of natural resource savings. *Resources, Conservation and Recycling,* Volume 54, pp. 229-234.
8. European Chemicals Agency, 2019. *How to comply with REACH Restriction 71, guideline for users of NMP (1-methyl-2-pyrrolidone),* s.l.: s.n.
9. Financial Times, 2019. *Congo, child labour and your electric car.* [Online] Available at: https://www.ft.com/content/c6909812-9ce4-11e9-9c06-a4640c9feebb [Accessed 19 November 2019].
10. Gaines, L., 2014. The future of automotive lithium-ion battery recycling: Charting a sustainable course. *Sustainable Materials and Technologies,* Volume 1-2, pp. 2-7.
11. Gaines, L., Richa, K. & Spangenberger, J., 2018. Key issues for Li-ion battery recycling. *MRS Energy & Sustainability,* Volume 5.
12. Goodenough, J., 2018. How we made the Li-ion rechargable battery. *Nature Electronics,* 1(204), p. 204.
13. Guerfi, A. et al., 2007. LiFePO4 water-soluble binder electrode for Li-ion batteries. *Journal of Power Sources,* Volume 163, pp. 1047-1052.
14. Gupta, U., 2019. *Innovation boosts lithium-ion battery recycling rate to over 80%.* [Online] Available at: https://www.pv-magazine.com/2019/03/25/innovation-boosts-lithium-ion-battery-recycling-rate-to-over-80/ [Accessed 14 November 2019].
15. Harper, G. et al., 2019. Recycling lithium-ion batteries from electric vehicles. *Nature,* Volume 575, pp. 75-86.
16. Hill, G., Heidrich, O., Creutzig, F. & Blythe, P., 2019. The role of electric vehicles in near-term mitigation pathways and achieving the UK's carbon budget. *Applied energy,* Volume 251.
17. Huang, B., Pan, Z., Su, X. & An, L., 2018. Recycling of lithium-ion batteries: Recent advances and perspectives. *Journal of Power Sources,* Volume 399, pp. 274-286.
18. Hunt, A., Matharu, A., King, A. & Clark, J., 2015. The importance of elemental sustainability and critical element recovery. *Green Chemistry,* 17(4), pp. 1949-1950.
19. IEA, 2019. *Global EV Outlook 2019,* s.l.: International Energy Agency.
20. IK4-CIDETEC, 2016. *Advanced Manufacturing Processes for Low Cost Greener Li-Ion Batteries*
21. Jiang, G. & Pickering, S., 2016. Recycling supercapacitors based on shredding and mild thermal treatment. *Waste Management,* Volume 48, pp. 465-470.
22. Joulié, M., Laucournet, R. & Billy, E., 2014. Hydrometallurgical process for the recovery of high value metals from spent lithium nickel cobalt aluminum oxide based lithium-ion batteries. *Journal of Power Sources,* Volume 247, pp. 551-555.
23. Kite-Powell, J., 2017. *This Recycling Robot Uses Artificial Intelligence To Sort Your Recyclables.* [Online] Available at: https://www.forbes.com/sites/jenniferhicks/2017/04/04/this-recycling-robot-

uses-artificial-intelligence-to-sort-your-recyclables/#7c9a95c62d35 [Accessed 19 November 2019].
24. Larcher, D. & Tarascon, J.-M., 2014. Towards greener and more sustainable batteries for electrical energy storage. *Nature Chemistry,* Volume 7, pp. 19-29.
25. Larsen, N., 2019. *Why Have Cobalt Prices Crashed.* [Online] Available at: https://internationalbanker.com/brokerage/why-have-cobalt-prices-crashed/ [Accessed 12 November 2019].
26. Loeffler, N. et al., 2017. Performance and Ageing Robustness of Graphite/NMC Pouch Prototypes Manufactured through Eco-Friendly Materials and Processes. *ChemSusChem,* Volume 10, pp. 3581-3587.
27. Lv, W. et al., 2018. A Critical Review and Analysis on the Recycling of Spent Lithium-Ion Batteries. *ACS Sustainable Chemistry & Engineering,* Volume 6, pp. 1504-1521.
28. McKinsey & Company, 2018. *Lithium and cobalt - a tale of two commodities,* s.l.: McKinsey & Company.
29. Melin, H. E., 2019. *Lithium batteries: Far more widely recycled than second-hand statistics imply.* [Online] Available at: https://www.energy-storage.news/news/batteries-all-active-materials-including-lithium-can-be-recycled-with-high [Accessed 25 November 2019].
30. Myung, S.-T.et al., 2017. Nickel-Rich Layered Cathode Materials for Automotive Lithium-Ion Batteries: Achievements and Perspectives. *ACS Energy Letters,* 2(1), pp. 196-223.
31. Oberhaus, D., 2019. *Tesla May Soon Have a Battery Which Can Last a Million Miles.* [Online] Available at: https://www.wired.com/story/tesla-may-soon-have-a-battery-that-can-last-a-million-miles/ [Accessed 23 November 2019].
32. Qiu, L. et al., 2014. Enhanced electrochemical properties of LFP cathode using the carboxymethylcellulose lithium (CMC-Li) as novel binder in lithium ion battery. *Carbohydrate Polymers,* Volume 111, pp. 588-591.
33. Richa, K., Babbit, C., Gaustad, G. & Wang, X., 2014. A future perspective on lithium-ion battery waste flows from electric vehicles. *Resources, Conservation and Recycling,* Volume 83, pp. 63-76.
34. Shi, Y. et al., 2018. Resolving the Compositional and Structural Defects of Degraded LiNixCoyMnzO2 Particles to Directly Regenerate High-Performance Lithium-Ion Battery Cathodes. *ACS Energy Lett.,* 3(7), pp. 1683-1692.
35. Shi, Y., Zhou, X. & Yu, G., 2017. Material and Structural Design of Novel Binder Systems for High-Energy, High-Power Lithium-Ion Batteries. *Accounts of Chemical Research,* Volume 50, pp. 2642 - 2652.
36. Sofiane, B. & Ravichandran, K., 2019. *Who will win the battery wars?.* [Online] Available at: https://www.greenbiz.com/article/who-will-win-battery-wars [Accessed 26 November 2019].
37. South China Morning Post, 2018. *China goes all out to secure lithum, cobalt supplies - key to dominating the world electric car market.* [Online] Available at: https://www.scmp.com/business/commodities/article/2149027/china-goes-all-out-secure-lithium-cobalt-supplies-key
38. Targray, 2017. *White Paper: Hydrophilic Binder Performance in Li-ion Batteries.* [Online] Available at: https://www.targray.com/articles/hydrophilic-binder [Accessed 18 November 2019].
39. US Department of the Interior, 2018. *Interior Seeks Public Comment on Draft List of 35 Minerals Deemed Critical to U.S. National Security and the Economy.* Washington: U.S. Government.
40. Wang, W.-Y., Yen, C. H., Lin, J.-L. & Xu, R.-B., 2018. Recovery of high-purity metallic cobalt from lithium nickel manganese cobalt oxide (NMC)-type Li-ion battery. *Journal of Material Cycles and Waste Management,* Volume 21, pp. 300-307.
41. Wang, X., Gaustad, G., Babbitt, C. & Richa, K., 2014. Economies of scale for future lithium-ion battery recycling infrastructure. *Resources, Conservation and Recycling,* Volume 83, pp. 53-62.
42. Winslow, K., Laux, S. & Townsend, T., 2018. A review on the growing concern and potential management strategies of waste lithium-ion batteries. *Resources, Conservation & Recycling,* Volume 129, pp. 263-277.
43. Wood, D. et al., 2018. Technical and economic analysis of solvent-based lithium-ion electrode drying with water and NMP. *Drying Technology,* 36(2), pp. 234-244.
44. Xiao, J., Li, J. & Xu, Z., 2017. Recycling metals from lithium ion battery by mechanical separation and vacuum metallurgy. *Journal of Hazardous Materials,* Volume 338, pp. 124-131.
45. Yao, Y. et al., 2018. Hydrometallurgical Processes for Recycling Spent Lithium-Ion Batteries: A Critical Review. *ACS Sustainable Chemistry & Engineering,* Volume 6, pp. 13611 - 13627.
46. Yun, L. et al., 2018. Metallurgical and mechanical methods for recycling of lithium-ion battery pack for electric vehicles. *Resources, Conservation & Recycling,* Volume 136, pp. 198-208.
47. Zeng, X., Li, J. L. & Singh, N., 2014. Recycling of Spent Lithium-Ion Battery: A Critical Review. *Environmental Science and Technology,* Volume 44, pp. 1129-1165.
48. Zhang, Z. et al., 2014. A comparative study of different binders and their effects on electrochemical properties of LiMn2O4 cathode in lithium ion batteries. *Journal of Power Sources,* Volume 247, pp. 1-8.
49. Zheng, R. et al., 2016. Optimized Li and Fe recovery from spent lithium-ion batteries via a solution-precipitation method. *Royal Society of Chemistry,* Volume 6, pp. 43613-43625.

Induction of Fatty Acid Synthesis in *Synechocystis* sp. PCC6803 by modification of culturing medium

Yen Yiu Ng and Yuen Yin Ho

Department of Chemical Engineering, Imperial College London, U.K.

ABSTRACT

Genetically modified cyanobacteria *Synechocystis* sp. PCC6803 PLY679 tailored for lipid accumulation displayed different growth characteristics and fatty acid quality under the exposure of phosphorus-limiting and carbon-enhancing environment. The strain was cultured in nine types of modified BG11 media in batch. The medium with 20% phosphorus reduction and 10% carbon enhancement showed a 14.4% increase in biomass content and a maximum fatty acid productivity of 2.73mg/L.day. When scaled-up to photobioreactor operation, a slower growth rate along with a lower lipid content by 18.5% were observed despite a greater lipid productivity of 6.6 mg/L.day. Overall, in contrast to biomass growth rate with variation in phosphorus concentration, fatty acid content did not show significant trends.

Keywords: *Synechocystis sp., phosphorus-limiting, carbon-enhancing, fatty acid, photobioreactor*

1. Introduction

Resolving the energy crisis has emerged as one of the top priorities worldwide as countries work towards securing their future energy supply. Increasing demand is placed upon the energy sector primarily due to the continual rise in world population and exponential economic growth of developing countries particularly China and India. In 2018, global energy demand increased by 2.3%, which is by far the fastest pace in this decade (IEA, 2018). In meeting this demand, current global energy relies heavily on the utilisation of fossil fuels including oil, coal and natural gas (BP, 2019). However, these reserves are predicted to deplete in the foreseeable future, and the combustion of fossil fuels emits greenhouse gases which lead to many negative environmental impacts, such as global warming due to release of CO_2.

In recent years, production of biofuels from microalgae and cyanobacteria has drawn considerable attention as a clean alternative to conventional fossil fuel (Milano et al., 2016). Microalgae are aquatic organisms capable of producing energy-rich compounds such as lipids and hydrocarbons through photosynthesis. Compared to other renewable resources such as solar, wind and hydropower, bioenergy has a greater potential since it is not limited by intermittency of the natural environment. As the third generation of biofuel production, microalgae is a more sustainable feedstock than agricultural crop since it is able to grow more efficiently with higher oil yields and does not compete with the food industry (Huang et al., 2017). Hence, developing large-scale cultivation of microalgae in the economical production of biofuels is of great interest in realising its promising future.

2. Background

There is a wide range of potential biofuels from products of microalgae, such as biohydrogen from molecular hydrogen (H_2), bioethanol from glucose and sucrose, and biomethane from fermentation of biopolymers. This study focuses on the ability of microalgae to synthesise fatty acids (FA) of carbon chain length of 14-20 which can be converted into biodiesel through esterification. The main challenges in commercialising this process is overcoming the biotechnical and economic obstacles within cultivation of the algae and harvesting of FA (Parma et al., 2011). Previous research had demonstrated the success in genetically modifying algal strains to allow secretion of free fatty acid (FFA) into the culture medium, hence providing a less energy intensive route in downstream harvesting than the traditional intracellular extraction methods (Liu, Sheng and Curtiss III, 2011). Irreversible removal of intracellular FFA would also theoretically promote the reaction, thereby increasing production of intracellular FFA. In addition, recent studies suggested the use of stress conditions such as nutrient limitation to increase lipid production of microalgae, however this often brings adverse effects on microalgal growth, thus hampering lipid productivity (Sun et al., 2018).

Synechocystis sp. PCC6803 is commonly used as a model cyanobacteria to understand microalgae photosynthetic pathways and fatty acid production since its genome is fully sequenced by previous extensive studies (Lu, 2010). Moreover, it is characterised by favourable properties such as the ability to survive in robust environment and flexible to genetic modifications (Yu et al., 2013). For cultivation of *Synechocystis* sp., the BG11 medium is used.

The key macronutrients required for growth of algae includes inorganic carbon (C), phosphorus (P) and nitrogen (N). As C accounts for 50% of the biomass dry weight (Ngyuen and Rittmann, 2016), it presents an excellent opportunity to increase C supply for a larger production of bioenergy and valuable products such as nutraceuticals or food supplements. However, problems have arisen in supplying CO_2 for industrial-scale biomass production due to its low carbon utilisation

efficiency and tendency to inhibit microalgal growth from the toxicity associated with the use of flue gas. Studies have shown that inorganic contaminants from coal-derived flue gas (As, Cd, Hg, etc.) negatively affected the biomass growth rate and lipid productivity, with an overall decrease of 67.5% and 77.9% respectively (Hess et al., 2017). Alternatively, HCO_3^--based cultivation turned out to be more superior than the conventional CO_2-based cultivation, with a higher fixation efficiency by 87.8% (Kim et al., 2017). Since the net HCO_3^- uptake of PCC6803 accounts for most of the photosynthetic rate in both low and high CO_2 conditions, the use of HCO_3^--based cultivation pronounces more benefits and can potentially compensate for the low carbon utilisation efficiency of the CO_2-based cultivation (Benschop, Badger and Dean Price, 2003).

Under carbon abundant situations, phosphorus was concluded to be the primary limiting factor which hampered growth of PCC6803 (Kim et al., 2010). Phosphorus plays an important role in metabolism as an essential component during energy transfer in the form of ATP, and is also involved in signal transduction and lipid synthesis. Liang (2012) experimented on *chlorella* sp. and reported an increase in lipid content and lipid productivity at low phosphorus concentration in BG11. Another study using *chlorella vulgaris* in BG11 medium ranked phosphorus concentration second after temperature in order of effects of cultivation parameters on lipid yield, and the maximum lipid yield was achieved at a lowered phosphorus concentration of 0.015g/L (Al-Rikabey and Al-Mayah, 2018). However, stress effects are highly dependent on the microalgae species. While there are several studies into effects of phosphorus concentration on growth of *Synechocystis* sp. (Fuszard et al., 2013; Kim et al., 2010), the simultaneous effects of phosphorus limitation on growth and FA production has not been studied previously.

On the other hand, studies on PCC6803 have reported an increase in both biomass productivity and lipid content under nitrogen deficient conditions by reducing $NaNO_3$ concentration in BG11 by 50% (Neag et al., 2019). Nitrogen as an imperative element in formation of proteins such as chlorophyll and enzymes, is a more widely studied parameter than phosphorus.

In this study, the genetically engineered *Synechocystis* sp. PCC6803 PLY679 was used to investigate the effect of medium composition on biomass growth and FFA production in batch cultures. In particular, microalgae performance under increased carbon availability and reduced phosphorus concentration in BG11 was evaluated. The optimal cultivation condition was then implemented in a scaled-up photobioreactor (PBR).

3. Methodology

3.1 Materials

9 types of modified BG11 medium were prepared based on the variation of K_2HPO_4 and Na_2CO_3 concentrations in the range of 0.024-0.036g/L and 0.022-0.028g/L respectively. The commercial BG11 medium recipe is provided in the Supplementary Information S.1. Antibiotic erythromycin (Sigma-Aldrich) was supplemented so as to ensure that only the genetically modified strain – *Synechocystis* sp. PCC6803 PLY679 grows in the medium.

N-hexane of 95% purity (Fischer Scientific) was used as a solvent for fatty acid extraction, followed by the addition of 37% hydrochloric acid (Fischer Scientific) in methanol (Sigma-Aldrich) for esterification.

3.2 Cyanobacteria and growth conditions

Synechocystis sp. PCC6803 PLY679 was isolated and maintained in the Research Laboratory, Department of Life Sciences at Imperial College London (Yunus and Jones, 2018). The genetically modified strain was pre-cultured in the commercial BG11 medium and inoculated in Nunc™ EasYFlask™ T-175 for 5-7 days. $NaNO_3$ concentration of the medium was reduced to 0.75g/L to limit nitrogen supply. Commercial (listed as **original**) and nitrate-limited media (**control**) were prepared in addition to the other samples to compare the growth under different nutrient conditions. All culturing media were supplemented with 20µg/mL antibiotic erythromycin and adjusted to pH 7 with NaOH and HCl to ensure favourable growth conditions of the cyanobacteria (Touloupakis et al., 2015).

100mL of cultures were grown in T-flasks as batch samples for 19 days and in 400mL flat-vessel PBRs for 10-15 days (FMT150, Photon Systems Instruments). For both experiments, the initial culture had an optical density of 0.2 at 730nm and the cultivation was carried out at 30°C with continuous illumination (white LED light) at 60µE/m^{2}s. Air enriched with 5% CO_2 (v/v) was continuously applied to the batch samples and were shaken at 150rpm in the incubator, whereas for the PBR no stirring was involved and a stream of 5% CO_2 (v/v) in air was directly injected to the culture at 5mL/min.

3.3 Analytical procedures

3.3.1 Biomass content and growth rates

Absorbance at 730nm were measured using Tecan Infinite® 200 PRO NanoQuant UV-Vis-NIR spectrophotometer. Microalgal growth was monitored every day by the determination of biomass content from the optical density at 730nm (OD_{730}). 10-fold dilution of the samples were done using deionised water before optical density measurements to keep the absorbance readings below 1. The biomass dry weight

concentrations (DW) indicated a linear positive correlation with OD_{730} (R^2 = 1), as shown in the Supplementary Information S.2, thus the productivity can be calculated as follows:

$$DW(g/L) = 0.3612 \times OD_{730} - 0.001 \quad (1)$$

The biomass production rate (μ) can then be determined by the following equation:

$$\mu\ (g/L.day) = \frac{DW_2 - DW_1}{t_2 - t_1} \quad (2)$$

where ($t_2 - t_1$) is the difference between the two time points.

3.3.2 Lipid extraction and esterification

Secreted FFAs were separated from the culture using the method described in Liu (2011) and Yunus (2018). Aliquots in Corning™ Pyrex™ round bottom threaded culture tubes were spiked with 20μL pentadecanoic acid as internal standard and acidified with 200μL 1M H_3PO_4, then extracted with 1.5mL of n-hexane. All samples were vortexed vigorously in between the addition of chemicals and centrifuged at 13000rpm for 5 minutes. The upper hexane layer was then transferred to a fresh glass tube and lipids were esterified at 65°C with 0.5μL of 1.25M HCl in methanol for 2 hours. When cooled to room temperature, 20μL of upper layer containing fatty acid methyl esters (FAMEs) was transferred to an insert in a gas chromatography (GC) vial.

3.3.3 Free fatty acid profile

Prior to analysing the composition and productivity of FFAs in samples, cultures were induced with $CoN_2O_6{\cdot}6H_2O$ to a concentration of 5μM after 48 hours of cultivation. 1mL aliquots were sampled out for esterification and off to GC analysis everyday thereafter.

FAMEs were subjected to gas chromatography/mass spectrometer (GC-MS) analysis using Aligent HP 6890/5973 GC-MS. The injector temperature was 250°C and helium at a flowrate of 1.5mL/min was used as the carrier gas. The heating programme used employed an oven temperature of 100°C for 30s, followed by an increase in heating rate of 30°C/min for 3 minutes. Each sample was analysed for 8 minutes and the intensity of the signal vs retention time was recorded.

4. Results and Discussion

4.1 Batch cultivation

Syneochocystis sp. PCC6803 PLY679 was cultured in 100mL shaked flask as batch samples for 19 days. The performance of the cyanobacteria in modified BG11 medium was analysed in terms of growth and FFA production in the following sections.

4.1.1 Biomass content

Growth of microalgae was quantified by converting OD_{730} measurements by the UV spectrophotometer into biomass content of the culture. Fig. 1 shows the effect of medium composition on biomass content on Day 19 of batch cultivation, values are represented as percentage difference relative to the control. The original medium showed a 3% decrease in biomass content compared to the nitrogen-limiting control sample. This verified the enhancement of biomass productivity when nitrogen concentration in BG11 was reduced to 0.75g/L, which Neag (2019) reported as the optimal nitrogen concentration for growth of PCC6803.

For cultures in the modified medium, all samples except at 40% phosphorus concentration reduction outperformed the control in terms of growth. The highest biomass content was obtained for *Synechocystis* sp. PCC6803 PLY679 cultured in modified BG11 medium with 20% P reduction and 10% C enhancement, which showed a 14.4% increase in biomass content relative to the control.

Compared to the control medium, phosphorus limitation in the modified BG11 medium lead to both a positive and negative effect on biomass content. A slight reduction in P up to 20% increased biomass content, meanwhile a 40% P reduction significantly decreased the biomass content. The lower biomass content for samples cultured in modified media of 40% P reduction reflected the importance of phosphorus for the growth of microalgae. Phosphorus, as an essential component in the production of chemical energy ATP, is utilised in reproduction of cells and various other cell activities. Therefore, the impairment in growth at higher reduction of P was as expected, and in accordance with similar studies on *Synechocystis* sp. (Fuszard et al., 2013).

A 20% reduction in P resulted in a higher biomass content than at 10% P reduction. The greatest difference was observed for cases at 10% carbon enhancement, in which a 20% P reduction showed a higher increase in biomass content relative to the control by 11.7%. The lower biomass content at 10% P reduction may be explained by the effect of 'luxury uptake' proposed by Wu (2012), which suggested that under excess nutrient supply, cells over-uptake phosphorus but did not lead to more growth. The increase in biomass content at 20% P reduction indicated that phosphorus supply is not under deprivation, however since phosphorus availability is less than at 10% P reduction, it is possible that more biomass with less phosphorous content is produced under a more limited P condition (Wu et al., 2012).

Variation in carbon supply of the modified BG11 medium gave contradicting results. 40% C enhancement showed the highest increase (7.8%) and lowest decrease (-7.3%) in biomass content relative to the control at 10% and 40% P reduction respectively. Meanwhile at 20% P reduction, 10% C enhancement resulted in the greatest increase in biomass content at 14.4% relative to the

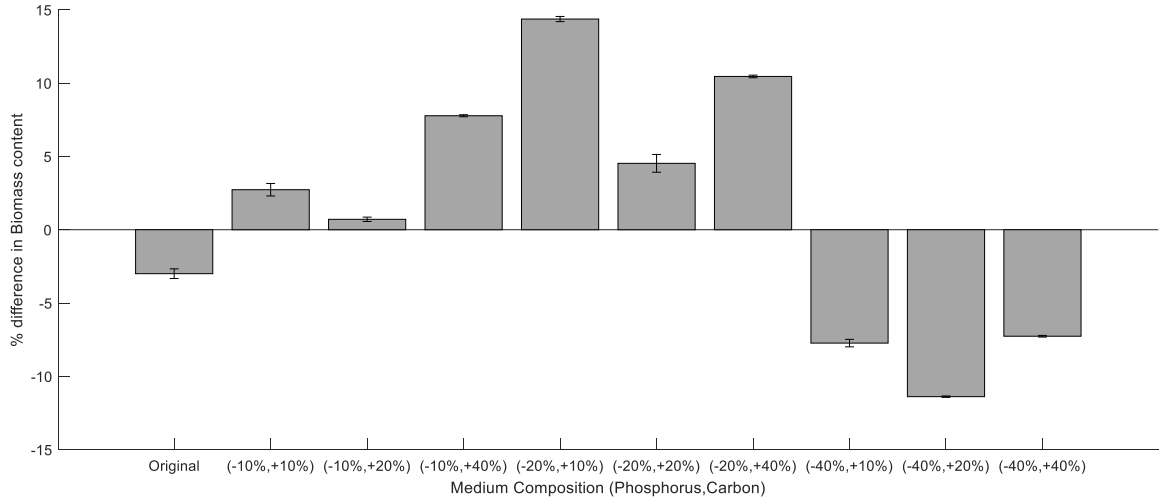


Fig. 1 *Effect of medium composition on biomass content on Day 19 of batch cultivation relative to control.*

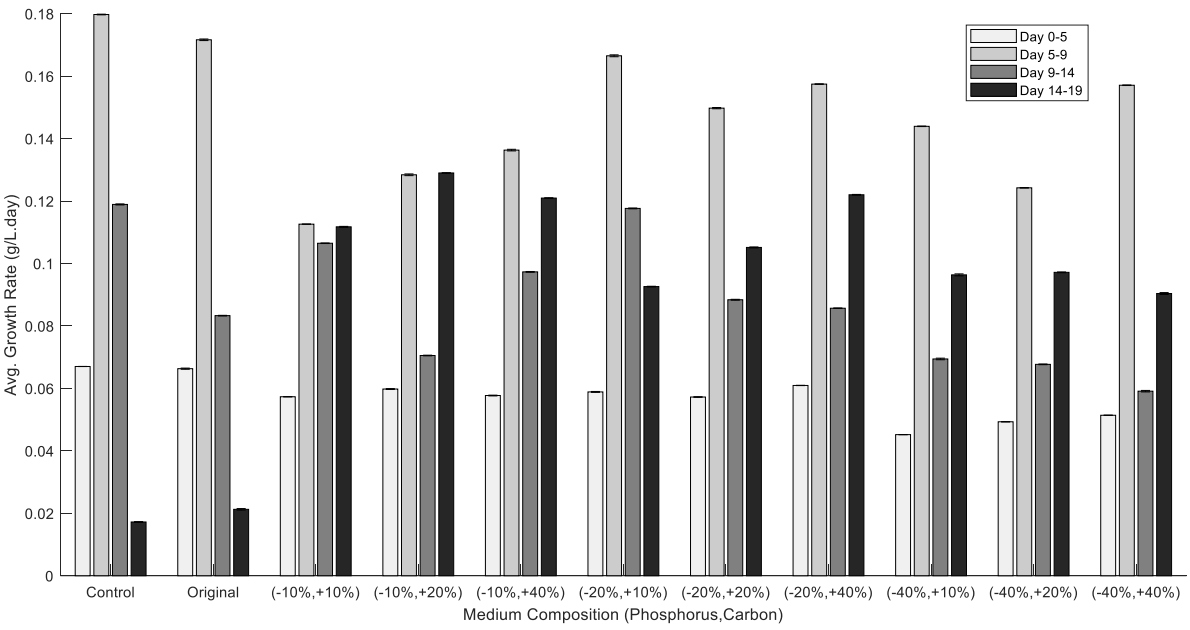


Fig. 2 *Effect of medium composition on biomass growth rates in batch cultivation.*

control. Hence, growth response of *Synechocystis* sp. to increasing C availability in the form of $NaHCO_3$ was not concentration-dependent in this experiment.

4.1.2 Growth rates

Monitoring growth rates throughout the cultivation period allows a better understanding of different growth phases. Cell growth can be divided into four distinct phases: lag, exponential, stationary and decline. Fig. 2 shows the effect of medium composition on average growth rates of batch culture at time intervals of 4-5 days.

All samples showed a maximum growth rate between Day 5-9, which indicated the exponential growth phase. The control sample displayed the highest maximum growth rate of 0.18g/L.day, as well as the highest growth rates between Day 0-5 and Day 9-14. In comparison to the control, a lower maximum growth rate was observed for cultures in the modified media, which is possibly due to the adjustment of microalgae to a different (P-limiting and C-enhancing) growth environment (Neag et al., 2019).

Among all modified media, maximum growth rate was the fastest at 0.17g/L.day for culture in the medium with 20% P reduction and 10% C enhancement. This medium also gave rise to the highest biomass content by Day 19 of batch cultivation from Section 4.1.1.

Between Day 14-19 of batch cultivation, there was a drastic reduction in growth rates for the control and original samples, which indicated signs of plateauing as biomass growth shifted towards the decline phase. Similarly, growth of *Synechocystis* sp. PCC6803 in BG11 medium was reported to enter the decline phase after 15 days of cultivation in Ma & Mi (2005). However, within this time period, growth rates for cultures in the modified medium remained relatively high ranging from 5.6-7.5 folds of the control. Other papers had reviewed adaptation responses such as

Table 1 Lipid content of the original and modified samples.

Fatty Acid	Control	Original	-10% P			-20% P			-40% P		
			+10% C	+20% C	+40% C	+10% C	+20% C	+40% C	+10% C	+20% C	+40% C
C14:0 (%)	10.2	16.2	19.7	19.7	17.6	15.9	17.5	15.3	13.5	15.9	22.3
C16:0 (%)	41.1	37.5	37.6	39.2	38.9	38.5	38.6	38.9	38.7	32.8	24.1
C18:0 (%)	48.7	46.3	42.7	41.1	43.5	45.6	43.9	45.7	47.8	51.3	53.6

regulation of gene translation and metabolic pathways of *Synechocystis* sp. under specific stress conditions (Sun et al., 2018; Herrero and Flores, 2008). Yet to fully validate this hypothesis as the reason for enhanced growth observed between Day 14-19 in the modified media, further studies to monitor changes in P and C concentration within the culture is required.

4.1.3 Free fatty acid profile

Analysis of FAMEs is essential to determine the potential of biofuel production from microalgae as composition of FAs is one of the key factors that determine the ignition quality and oxidative stability of the biofuels (Lee et al., 2010). FA profiles consisting primarily palmitic acid (C16:0) and stearic acid (C18:0) are perfect candidates for the production since short-chain saturated FAs have a higher Cetane number and a better oxidative stability (Benjumea et al., 2011).

The relative contents of FFAs were calculated based on the peak area from chromatograms and the profile of *Synechocystis* sp. PCC6803 PLY679 is shown in Table 1. According to Merritt (1991), it was reported that under the same cultivation conditions, PCC6803 contained large traces of polyunsaturated linoleic acid (C18:2) and α-linolenic acid (C18:3), with C16:0 as the predominant species. In contrast, polyunsaturated species were not detected in the investigated genetically-modified strain and C18:0 accounted for most of the FA composition, with myristic acid (C14:0) being the lowest in content. The proportion of C18:0 was consistently greater than other FAs, regardless of the variation in P and C concentrations.

An overall increase in lipid content was observed when only nitrogen concentration was limited in the medium. Apart from the three media with a 40% decrease in P, the C16:0 content was not significantly affected as it accounted for approximately 38.5% of the fatty acid composition in all other modified media. However, these media accommodated a lower C18:0 proportion than both the control and original samples. This is not in agreement with studies of the same species reported by Rodolfi (2009), in which a decrease in biomass productivity and increase in C16:0 acid were observed in P deprivation environment. As mentioned in Section 4.1.2, growth of microalgae in 10% and 20% P reduction media did not undergo nutrient deprivation. Hence, lipids produced from these samples were not as rich as expected. An in-depth analysis of samples with

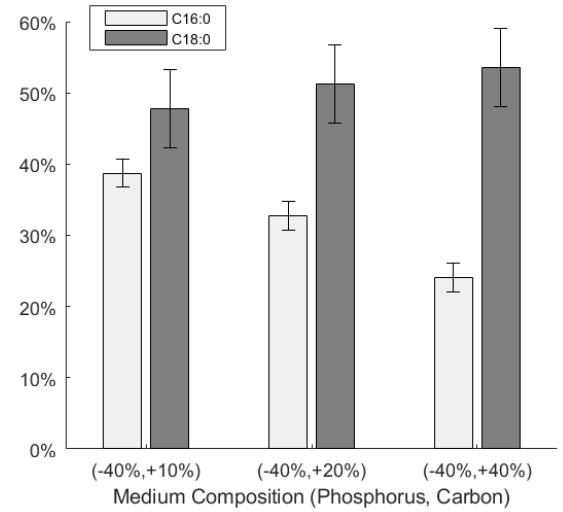


Fig. 3 Lipid content with increasing C concentration at 40% P reduction.

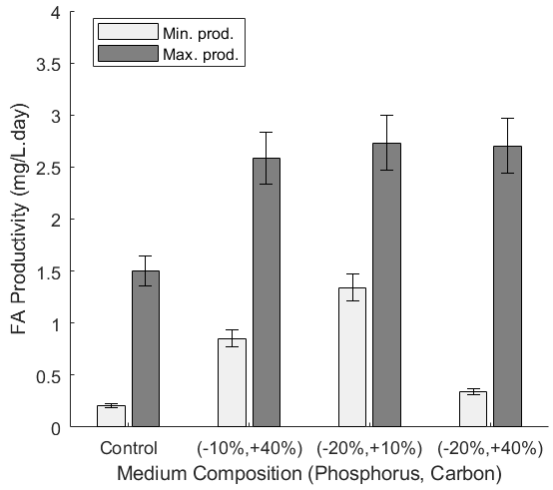


Fig. 4 Max. and min. lipid productivity of different media.

40% reduction in P as shown in Fig. 3 indicated an overall reduction in the lipid content as the C concentration in the medium increases. Drastic changes in the C16:C18 acid ratio were also observed, in which C16:0 % dropped to a minimum of 24.1% and C18:0 % increased to a maximum of 53.6%.

From studies by Rodolfi (2009), where nutrient stress was introduced to the culture under appropriate exposure to light, photosynthesis resumed at a slower rate and the flow of fixed carbon diverted from protein to lipid synthesis. Generally, a greater limitation in nutrient concentration of the medium reduces overall biomass productivity but results in a richer lipid content.

Comparing Fig. 1 and Fig. 3, it was recognised that the culture in 40% P reduction media had the lowest productivity among all samples. Similar to its overall FA content, it was also affected by the increase in C concentration at P-limiting conditions. Hence, medium with a 40% P reduction are not the most favourable culturing.

4.1.4 Summary

When selecting a medium for optimal cultivation of PCC6803 PLY679, achieving a richer lipid content is one of the main goals as this improves the efficiency of extraction, yet biomass productivity is equally important. Referring to both the overall biomass content that PCC6803 can produce in the modified media and its lipid productivity shown in Fig. 1 and Fig. 4 respectively, the strain showed the best performance in BG11 medium with 20% in P reduction and 10% in C enhancement as it displayed its highest potential in both maximum and minimum production rate of 2.73mg/L.day and 1.34mg/L.day respectively. Thus (−20%P, +10%C) was selected as the medium for the PBR operation in comparison with the control medium.

4.2 Scale-up cultivation

Selection of cultivation system is usually based on factors such as effectiveness and cost. Open raceway pond systems are one of the most commonly used type as they give a larger scale biomass production at a lower capital cost (Narala et al., 2016). However, the production can be easily altered by external factors such as seasonal variations and contamination by unwanted organisms. The cyanobacteria used in this experiment is genetically modified to maximise the lipid accumulation and since antibiotic is used specifically for this strain, environmental concerns will be raised for open ponds as it can cause ecological imbalances in the natural communities (Beacham, Sweet and Allen, 2017).

Alternatively, production in closed systems like PBRs provide adequate supply (light, CO_2-enriched air, buffer) to the culture to ensure microalgae is at its optimal production throughout the entire culturing process. Even though PBRs are more expensive, they give a prolonged and more controllable cultivation. Hence, a bench-scale PBR was implemented for the scale-up cultivation.

Synechocystis sp. PCC6803 PLY679 was cultured in the PBR for 10 days using the nitrogen-limited control BG11 medium, and 15 days using the modified BG11 medium with 20% P reduction and 10% C enhancement. Comparison between the two types of cultivation modes of PBR and batch regarding growth and FFA production of the cyanobacteria were discussed in the following sections.

4.2.1 Comparison of biomass content and growth rates between PBR and batch cultivation

The PBR offers a more precise control of the microalgal growth environment in terms of light intensity and temperature compared to batch cultivation performed in the incubator. For growth in the control medium, the batch mode achieved a biomass content of 1.1261g/L by Day 10 of cultivation as presented in Table 2, which is higher than that in the PBR by a third. On the other hand, similar biomass contents were observed in both modes of cultivation for growth in the modified medium with 20% P reduction and 10% C enhancement, at 1.7095g/L in batch and only marginally lower in the PBR by Day 15 of cultivation.

Previously in Section 4.1.2, Day 5-9 was identified as the time period with maximum growth rates for this strain when cultured in batch. For both media, there was a significantly higher growth rate between Day 5-9 of cultivation in the batch mode in contrast to the PBR as shown in Fig. 5. In fact, growth rates in the PBR between Day 5-9 was reduced by half for the control and by 28% for the (−20%P, +10%C) medium. Growth conditions were kept constant between the batch and PBR, but the most significant difference was the absence of stirring in the PBR. Hence, decrease in maximum growth rates under PBR operation was most probably due to insufficient mixing, which prevented efficient transfer of materials from the medium to cells. This reflected the importance of mixing in PBR, which

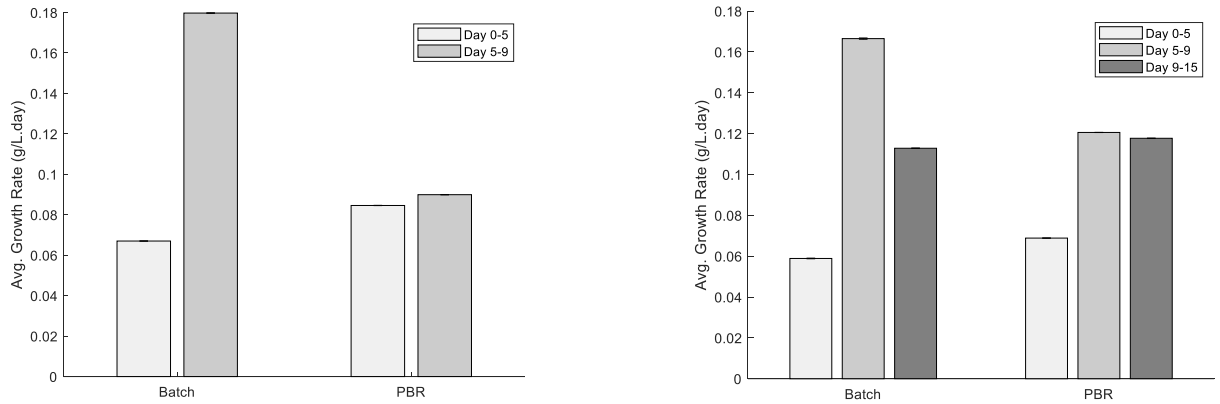


Fig. 5 Comparison of average growth rates between batch and PBR modes of cultivation for microalgal growth using the control medium (Left) and modified BG11 medium with 20% P reduction and 10% C enhancement (Right).

was highlighted in Ugwu (2008) to induce turbulent flows and distribute nutrients, thus allowing optimal growth.

Table 2 Comparison of biomass content obtained by the end of cultivation period (Day 15 for control medium and Day 10 for (-20%P, +10%C) medium) between batch and PBR modes.

	Biomass content (g/L)	
	Control medium	**(-20%P, +10%C) medium**
Batch	1.1261 ± 0.0003	1.7095 ± 0.0002
PBR	0.7890 ± 0.0003	1.6566 ± 0.0003

4.2.2 Comparison of free fatty acid profiles between PBR and batch cultivation

Lipid productivity was greatly enhanced for all media when cultivation was scaled-up according to Fig. 6, especially for sample (−20%P, +10%C), which a maximum rate of 6.6mg/L.day was recorded. However, a drop in overall lipid content of 18.5% was observed according to Fig. 7. Similarly for the control, a significant 20.5% decrease in overall lipid content was observed, and the C16:0 acid was more dominant than C18:0.

Possible reasons for the unexpected drop in lipid content can be insufficient mixing as mentioned in Section 4.2.1, causing an uneven light intensity inside the PBR. Microalgal growth in PBR would likely undergo flocculation and biofilm formation during long-term cultivation, leading to a reduction in light incident to certain parts of the reactor. This in turn reduces the cyanobacteria's phototrophic efficiency, which decreases the biomass productivity per unit of light and thus negatively impacting lipid accumulation (Yu et al., 2013; Xu et al., 2009). Since only 5mL/min of CO_2-enriched air was injected to the culture, aeration rate is not high enough for liquid circulation. Floccules formed at the bottom of the reactor would also clog the gas supply, stopping effective gas-liquid mass transfer between CO_2 and the microalgae. If mixing is incorporated, it would prevent oxygen accumulation in some regions of the reactor and provide better mass transfer to boost the cell growth rate, leading to greater amount of fatty acids produced.

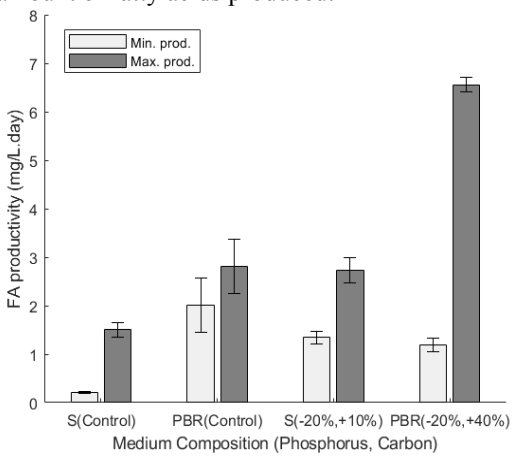


Fig. 6 Lipid productivity of different modes of cultivation (B: batch, PBR: photobioreactor).

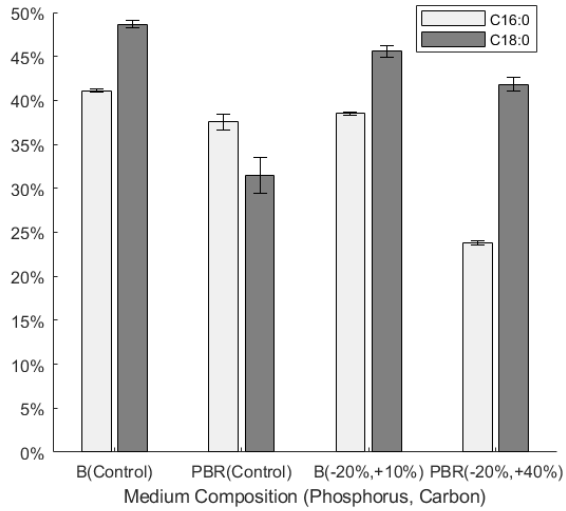


Fig. 7 Lipid content at different modes of cultivation.

5. Conclusion

This study illustrated the effect of varying nutrient concentration in BG11 medium for cultivation of *Synechocystis* sp. PCC6803 PLY679. Modified BG11 medium with 20% phosphorus reduction and 10% carbon enhancement (corresponding to 0.032g/L K_2HPO_4 and 0.022g/L Na_2CO_3) showed the best performance in optimising both biomass and FA productivity under batch cultivation. Slight reduction in phosphorus concentration up to 20% was favourable, resulting in increased biomass content and FA productivity, and did not alter the proportion of C16:0 and C18:0 acid produced. Greater reduction in P by 40% was proven to be an unfavourable environment, hence should be avoided in cultivation for this strain. Due to inconclusive results in trends on biomass content and FA composition with increasing C concentration, it is suggested not to waste valuable resources by increasing C supply during cultivation. Upon scale-up of the cultivation process, although the PBR gave a more consistent biomass growth rate and a greater lipid productivity, the conditions used on batch samples were not applicable on the PBR cultivation as it is limited by other factors besides the lack of nutrients.

6. Outlook

To optimise both the biomass growth rate and lipid productivity, a two-stage hybrid cultivation involving a nutrient-sufficient medium to allow maximum cell growth, followed by a nutrient-deficient stage to induce lipid production can be implemented (Ge et al., 2016). A two-phase separate system benefits the FA production without having serious concerns on contamination (Narala et al., 2016).

Choice of solvent is crucial for an effective lipid extraction as Sheng (2011) reported that at least 20% of lipid was lost when hexane was used for extraction, whereas chloroform and methanol-based solvent used in

Bligh and Dyer and Folch methods displayed the greatest lipid recoveries. However, the toxicity associated with methanol and chloroform raise environmental concerns and are considered unfavourable for large-scale extraction (U.S. EPA, 2009). Thus, methanol can be supplemented with MTBE as an alternative solvent for lipid extraction.

Acknowledgements

Authors would like to express their gratitude to Mr Nuttapon Vachiraroj, Mr Ian Sofian Yunus and Dr Fessehaye Zemichael for their support and guidance in this research.

References

Al-Rikabey, M. N. and Al-Mayah, A.M. (2018) Cultivation of Chlorella Vulgaris in BG-11 media using Taguvhi method. *Journal of Adv Research in Dynamical & Control system*, pp.19-30

Beacham, T., Sweet, J. and Allen, M. (2017). Large scale cultivation of genetically modified microalgae: A new era for environmental risk assessment. *Algal Research*, 25, pp.90-100.

Benjumea, P., Agudelo, J. and Agudelo, A. (2011). Effect of the Degree of Unsaturation of Biodiesel Fuels on Engine Performance, Combustion Characteristics, and Emissions. *Energy & Fuels*, 25(1), pp.77-85.

Benschop, J., Badger, M. and Dean Price, G. (2003). *Photosynthesis Research*, 77(2/3), pp.117-126.

BP p.l.c. (2019). *BP Statistical Review of World Energy 2019*. London: BP p.l.c.

Fuszard, M., Ow, S., Gan, C., Noirel, J., Ternan, N., McMullan, G., Biggs, C., Reardon, K. and Wright, P. (2013). The quantitative proteomic response of *Synechocystis* sp. PCC 6803 to phosphate acclimation. *Aquatic Biosystems*, 9(1), p.5.

Ge, S., Champagne, P., Plaxton, W., Leite, G. and Marazzi, F. (2016). Microalgal cultivation with waste streams and metabolic constraints to triacylglycerides accumulation for biofuel production. *Biofuels, Bioproducts and Biorefining*, 11(2), pp.325-343.

Herrero, A. and Flores, E. (2008). *The cyanobacteria*. Norfolk, U.K.: Caister Academic Press, p.117.

Hess, D., Napan, K., McNeil, B., Torres, E., Guy, T., McLean, J. and Quinn, J. (2017). Quantification of effects of flue gas derived inorganic contaminants on microalgae growth system and end fate of contaminants. *Algal Research*, 25, pp.68-75.

Huang, Y., Lai, C., Wu, B., Lin, K., Wu, J.C.S., Hossain, M.S.A., Yamauchi, Y., Wu, K.C. (2017). Advances in bioconversion of microalgae with high biomass and lipid productivity. *Journal of the Taiwan Institute of Chemical Engineers*, 79, pp.37-42

Iea.org. (2019). Global energy demand rose by 2.3% in 2018, its fastest pace in the last decade. [online] Available at: https://www.iea.org/newsroom/news/2019/march/global-energy-demand-rose-by-23-in-2018-its-fastest-pace-in-the-last-decade.html [Accessed 25 Nov 2019].

Kim, G., Heo, J., Kim, H. and Han, J. (2017). Bicarbonate-based cultivation of Dunaliella salina for enhancing carbon utilization efficiency. Bioresource Technology, 237, pp.72-77.

Lee, J., Yoo, C., Jun, S., Ahn, C. and Oh, H. (2010). Comparison of several methods for effective lipid extraction from microalgae. *Bioresource Technology*, 101(1), pp.S75-S77.

Liu, X., Sheng, J. and Curtiss III, R. (2011). Fatty acid production in genetically modified cyanobacteria. *Proceedings of the National Academy of Sciences*, 108(17), pp.6899-6904.

Lu, X. (2010). A perspective: Photosynthetic production of fatty acid-based biofuels in genetically engineered cyanobacteria. *Biotechnology Advances*, 28(6), pp.742-746.

Ma, W. and Mi, H. (2005). Expression and activity of type 1 NAD(P)H dehydrogenase at different growth phases of the cyanobacterium, *Synechocystis* PCC 6803. *Physiologia Plantarum*, 125(1), pp.135-140.

Merritt, M., Rosenstein, S., Rachel, C., Chou, H. and Allen, M. (1991). A comparison of the major lipid classes and fatty acid composition of marine unicellular cyanobacteria with freshwater species. *Archives of Microbiology*, 155(2), pp.107-113.

Milano, J., Ong, H., Masjuki, H., Chong, W., Lam, M., Loh, P. and Vellayan, V. (2016). Microalgae biofuels as an alternative to fossil fuel for power generation. *Renewable and Sustainable Energy Reviews*, 58, pp.180-197.

Musa, M., Ayoko, G.A., Ward, A., Rosch, C., Brown, R.J., Rainey, T.J. (2019) Factors affecting microalgae production for biofuels and the potential of chemometric methods in assessing and optimising productivity. *Cells*, 8, pp.851

Narala, R., Garg, S., Sharma, K., Thomas-Hall, S., Deme, M., Li, Y. and Schenk, P. (2016). Comparison of Microalgae Cultivation in Photobioreactor, Open Raceway Pond, and a Two-Stage Hybrid System. *Frontiers in Energy Research*, 4.

Neag, E., Török, A., Cadar, O., Băbălău – Fuss, V. and Roman, C. (2019). Enhancing lipid production of *Synechocystis* PCC 6803 for biofuels production, through environmental stress exposure. *Renewable Energy*, 143, pp.243-251.

Nguyen, B. and Rittmann, B. (2016). Effects of inorganic carbon and pH on growth kinetics

of *Synechocystis* sp. PCC 6803. *Algal Research*, 19, pp.363-369.

Parmar, A., Singh, N., Pandey, A., Gnansounou, E. and Madamwar, D. (2011). Cyanobacteria and microalgae: A positive prospect for biofuels. *Bioresource Technology*, 102(22), pp.10163-10172.

Rodolfi, L., Chini Zittelli, G., Bassi, N., Padovani, G., Biondi, N., Bonini, G. and Tredici, M. (2009). Microalgae for oil: Strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnology and Bioengineering*, 102(1), pp.100-112.

Sheng, J., Vannela, R. and Rittmann, B. (2011). Evaluation of methods to extract and quantify lipids from *Synechocystis* PCC 6803. Bioresource Technology, 102(2), pp.1697-1703.

Sun, X., Ren, L., Zhao, Q., Ji, X. and Huang, H. (2018). Microalgae for the production of lipid and carotenoids: a review with focus on stress regulation and adaptation. *Biotechnology for Biofuels*, 11(1).

Touloupakis, E., Cicchi, B., Benavides, A. and Torzillo, G. (2015). Effect of high pH on growth of *Synechocystis* sp. PCC 6803 cultures and their contamination by golden algae (Poterioochromonas sp.). Applied Microbiology and Biotechnology, 100(3), pp.1333-1341.

Ugwu, C., Aoyagi, H. and Uchiyama, H. (2008). Photobioreactors for mass cultivation of algae. *Bioresource Technology*, 99(10), pp.4021-4028.

US EPA. IRIS Toxicological Review of Methanol (External Draft). US Environmental Protection Agency, Washington, DC, EPA/635/R-09/, 013, 2009.

Wu, Y., Yu, Y., Li, X., Hu, H. and Su, Z. (2012). Biomass production of a Scenedesmus sp. under phosphorous-starvation cultivation condition. *Bioresource Technology*, 112, pp.193-198.

Xu, L., Weathers, P., Xiong, X. and Liu, C. (2009). Microalgal bioreactors: Challenges and opportunities. *Engineering in Life Sciences*, 9(3), pp.178-189.

Yu, Y., You, L., Liu, D., Hollinshead, W., Tang, Y. and Zhang, F. (2013). Development of *Synechocystis* sp. PCC 6803 as a Phototrophic Cell Factory. Marine Drugs, 11(8), pp.2894-2916.

Yu, Y., You, L., Liu, D., Hollinshead, W., Tang, Y. and Zhang, F. (2013). Development of *Synechocystis* sp. PCC 6803 as a Phototrophic Cell Factory. *Marine Drugs*, 11(8), pp.2894-2916.

Yunus, I. and Jones, P. (2018). Photosynthesis-dependent biosynthesis of medium chain-length fatty acids and alcohols. Metabolic Engineering, 49, pp.59-68.

The oxidation of Printex-U in NO_2 at various temperatures and concentrations.

Khilan Shah and Ramya Gocouladasse
Supervisor: Dr C.J. Tighe
Department of Chemical Engineering, Imperial College London, U.K.

Abstract The kinetics of the oxidation of Printex-U by NO_2 have been measured in a packed bed at a range of temperatures (400-500°C) and concentrations of NO_2 (500-800ppm). The reactions were carried out in a packed bed reactor, where the composition of the outlet flow was analysed by a Fourier-transform infrared spectroscope. Sources of error were identified with the existing apparatus, and modifications made, particularly with setting up reactant flow via a bypass valve. The shrinking core model and pseudo-homogenous model were compared to identify which best represents the kinetics of the oxidation reaction. The former was found to model oxidation well over a larger range of conversion, but was limited by the assumption that the particles are uniform in diameter. The reaction was found to be approximately first order with an apparent activation energy of 83 $kJ\ mol^{-1}$. Mass balances on carbon, oxygen and nitrogen showed that only CO, CO_2 and NO were produced during the oxidation of Printex-U with NO_2 .The assumptions that the particles of soot are isothermal and experience no diffusional limitations were also evaluated.

Introduction

The oxidation of soot has been an important area of study over the past few decades, especially since the proliferation of diesel cars - diesel engines accounted for over 17% of carbon black emissions in 2000. It is important to reduce soot particulate emissions since they can have deleterious impacts on human health, having been associated with smog and carcinogenic effects in humans. A better understanding of the kinetics of soot oxidation can help improve the operation of diesel particulate filters, especially during its regeneration cycle where the carbon caught in the filter is oxidised from the engine exhaust gases. This can help reduce the particulate matter released in newer diesel cars. Nitrogen dioxide is found within the exhaust of diesel engines and has been found to significantly increase the rate of carbon oxidation. The two main reactions occurring during NO_2 oxidation are:

$$2NO_{2\,(g)} + C_{(g)} \rightarrow CO_{2\,(g)} + 2NO_{(g)}$$
$$NO_{2\,(g)} + C_{(g)} \rightarrow CO_{(g)} + NO_{(g)}$$

Activation energies for Printex-U in oxygen have been determined as $\sim 168\ kJmol^{-1}$, in contrast to a much lower activation energy in nitrogen dioxide in the range of $50 - 80\ kJmol^{-1}$. A lower activation energy would allow oxidation to occur at significantly lower temperatures, potentially lowering the required temperatures during regeneration of particulate filters.

Reaction regimes

It has been suggested[1] that the 'intrinsic' rate typically measured in oxidation reactions are actually representative of three regimes, differentiated by their rate limiting steps:

- Regime 1: the concentration of the reactant is zero on the particle surface, so oxidation rate is solely determined by external film diffusion of the reactant.
- Regime 2: The oxidation rate is limited by both the reactant diffusing through the particle pores and the kinetic rate. This regime is characterised by a concentration gradient inside the particle.
- Regime 3: The oxidation rate is limited only by the kinetic rate as the reaction is not subject to any diffusional limitations, the concentration of reactant is the same in all regions.

Reaction rates measured in regime 3 are considered to most accurately represent the intrinsic rate of the oxidation reaction. For this reason, previous studies, have been carried out by mixing a small, *ex situ*, sample of soot in an inert material such as quartz[2] or silicon carbide[3], amongst others, in order to reduce mass transfer limitations.

Measuring kinetics

For low temperature oxidation reactions, previous experiments[4] have utilised Thermal Gravimetric Analysis (TGA) to study carbon/oxygen reactivity. Here, the sample of soot is heated to the required temperature in an inert gas, the reactant (oxygen) is then introduced and the change in reactant mass during oxidation is measured.

It is noted that reactions can be operated under kinetic control, both isothermally and non-isothermally – the main advantage of this method is that the quantity of adsorbed material such as organics can be determined, and the temperature where the reaction first occurs can be found in temperature ramping experiments. Experiments have also been carried out in crucibles[5], where the relatively larger surface area of the soot per unit volume of the sample, coupled with diffusional limitations of the oxidant between the mouth of the crucible and the surface of the bed, meant mass transfer limitations had to be considered.

Experiments are also typically carried out in fixed bed reactors[3], operated isothermally to study kinetic measurements. A bed containing a diluted soot sample is initially heat treated with an inert gas prior to oxidation. A dilute soot sample is used to limit the change in concentration of reactant and maintain differential operation[28].

Morphology of soot

Particles of soot typically exist as agglomerates[6] (shown in Figure 1), $\sim 100\mu m$ in size, consisting of loose, porous particles $\sim 0.1 - 1\mu m$ in size which are made up of smaller carbonaceous spherules, $\sim 10 - 50nm$ in size, that are clustered together along with trace amounts of

adsorbed and vapour phase hydrocarbons, sulphates, and soluble organic fraction (SOFs).

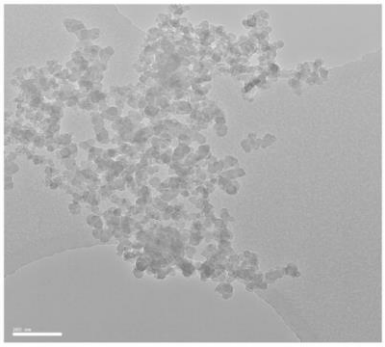


Figure 1: Transmission Electron Microscopic (TEM) image of Printex-U showing an agglomerate of particles $< 50nm$ in size. [33]

Previous studies have concluded that during oxidation reactions, the reactant diffuses into the particle to react with carbon spherules. The specific surface area of the sample was found to increase in a previous experiment[7], where agglomerates of soot were oxidised by O_2 and changes in the microstructure studied with an analytical electron microscope. It was found that as SOFs were released from the sample and the reaction progressed, the specific area of the sample increased – i.e. the particle became more porous during the reaction as previously blocked pores became accessible by reactants. The increase in specific area of the particle, S_p, can be written as Eq. 1 for an γth order model; $S_{p,0}$ is the initial specific surface area, X_C is the conversion of carbon and γ is the reaction order with respect to carbon[1]. For a shrinking particle with a spherical geometry, $\gamma = \frac{2}{3}$[1].

$$S_p = S_{p,0}(1 - X_C)^{\gamma} \quad (1)$$

It has also been discussed[2] that due to the porosity of the particle being sufficiently large, and the size of the particle itself being sufficiently small, diffusional limitations can be neglected (i.e. effectiveness factor, $\eta = 1$), so representing the soot as a porous particle, or as individual spherules that burn separately, are equivalent.

Proposed mechanisms of the oxidation of soot in NO_2

The two reactive sites on soot particles typically cited in literature are the basal sites and edge-site carbons, with the latter being more reactive with NO_2 [8]. It is also well studied[6,9], that the oxidation rate of soot in NO_2 is much higher than in O_2 alone, especially at high temperatures, due to the relative ease of dissociation of NO_2 on edge-site carbons, CC_f, compared to O_2; the former has a bond dissociation enthalpy[9] of 498 $kJ\ mol^{-1}$ and the latter, 306 $kJ\ mol^{-1}$.

A mechanism for the NO_2 preferentially (if not only) oxidising edge site carbons[8] has been proposed via the following reactions[9]:

$$CC_f + NO_2 \rightarrow CC_f(O) + NO$$

$$CC_f(O) + NO_2 \rightarrow C(O)C_f(O) + NO$$

Where $CC_f(O)$ and $C(O)C_f(O)$ are carbon atoms that have been oxidised on the surface and can decompose to CO or CO_2.

Considering all of this, the aims of this study were to further explore the oxidation of carbon in nitrogen dioxide by:

- Verifying common assumptions such as neglecting mass transfer limitations and isothermal operation, made in this study.
- Minimising sources of error in the experimental apparatus used in previous research.

Pseudo-homogeneous model

The pseudo-homogeneous model describes uniform oxidation throughout a porous particle. By assuming that the concentration of oxidant within the particle is the same as that in the bulk stream[10], the heterogeneous gas-solid oxidation reaction can be approximated as a homogenous reaction, taking place at all points both on the surface and within the spherule[11].

The rate of oxidation of one of these spherules can then be outlined as in Eq. 2, applicable to reactions where fine solid particles (e.g. Printex-U) are embedded in an inert material (e.g. quartz sand), and assuming the reaction rate is independent of the concentration of carbon[10]. Here, N_C is the number of moles of carbon contained in the spherule in mol, k_S is the rate constant in terms of the reacting surface area of the particle in $mol^{3n-2}mol^{1-n}s^{-1}$, $C_{NO_2}^n$ is the concentration of NO_2 in $mol\ m^{-3}$, n represents the order of the reaction kinetics, S_p is the surface area per unit volume of the particle and in $m^2\ m^{-3}$ and V_P is the volume of the spherule in m^3.

$$-\frac{dN_c}{dt} = k_s C_{NO_2}^n S_p V_p \quad (2)$$

It is also assumed that, during the oxidation reaction, the volume of the particle remains constant at its starting value, $V_{p,o}$. The moles of carbon present in the particle can be expressed, as in Eq. 3, as a function of the conversion of carbon, X_C, the constant molar density of soot, ρ_M in $mol\ m^{-3}$, the porosity of the spherule ε_p, which increases from an initial value $\varepsilon_{p,0}$, to 1 when all the carbon has been oxidised, and $V_{p,0}$.

$$N_C = (1 - X_C)\rho_M(1 - \varepsilon_{p,0})V_{p,0} \quad (3)$$

The specific surface area of the particle can be assumed to follow the trend described in Eq. 1; for uniform burning, γ is equivalent to unity. Substituting Eq. 1 and 3 into Eq. 2 gives:

$$-\rho_M(1 - \varepsilon_{p,0})V_{p,o}\frac{d(1-X_C)}{dt} = k_s C_{NO_2}^n S_{p,0}(1 - X_C)V_{p,0} \quad (4)$$

Rearranging Eq. XX and integrating with respect to time:

$$\ln(1 - X_C) = -\frac{k_s S_{p,0}}{\rho_M(1-\varepsilon_{p,0})} C_{NO_2}^n t \quad (5)$$

A volumetric rate constant, k_v, can then be defined and used to truncate Eq. 5 in Eq. 7:

$$k_v = \frac{k_s S_{p,0}}{\rho_M(1-\varepsilon_{p,0})} \quad (6)$$

$$\ln(1 - X_C) = -k_v C_{NO_2}^n t \quad (7)$$

Shrinking core model

The predominant assumption associated with the shrinking core model is that the carbon spherules are non-porous and so the oxidation reaction is strictly limited to the surface. The equation utilises the rate equation for the shrinking core model[2] for spherical particles of unchanging size, and assumes kinetic control by removing all diffusional limitations.

Since the particle is non-porous, Eq. 3 can be adapted to remove the porosity term and constant volume assumption into Eq. 8.

$$N_C = (1 - X_C)\rho_M V_p \qquad (8)$$

Exploiting the definition of the specific surface area as $S_{p,0} = \frac{\pi d_{p,0}^2}{\pi \frac{d_{p,0}^3}{6}} = \frac{6}{d_{p,0}}$, where d_p is the diameter of the particle, and substituting this result and Eq. 8 into Eq. 2 yields:

$$-\rho_M \frac{d(1-X_C)}{dt} = \frac{6k_s}{d_{p,0}} C_{NO_2}^n (1 - X_C)^{\frac{2}{3}} \qquad (9)$$

Where $\gamma = \frac{2}{3}$ for a sphere. However, the Printex-U sample in the packed bed contains spherules with a range of diameters not currently accounted for. An earlier report concluded[2] that the gradient of a plot of $1 - (1 - X_C)^{\frac{1}{3}}$ against time is unaffected by a distribution in $d_{p,0}$ when the log mean of the diameters, $(d_{p,0})_{lm}$ was considered and $X_C < 0.65$. Substituting this result into Eq. 9 and integrating with respect to time results in Eq. 10.

$$1 - (1 - X_C)^{\frac{1}{3}} = \frac{2k_s}{d_{p,0}\rho_M} C_{NO_2}^n \qquad (10)$$

Defining a logarithmic mean volumetric rate constant, $(k_v)_{lm}$ as in Eq. 11, allows for further simplification of the rate equation to Eq. 12, which uses a log mean distribution of $\sigma_{lm} = 0.4$[2].

$$(k_V)_{lm} = \frac{2k_s}{(d_{p,0})_{lm}\rho_M} \qquad (11)$$

$$1 - (1 - X_C)^{\frac{1}{3}} = (k_V)_{lm} C_{NO_2}^n t \qquad (12)$$

Comparison of models

The volumetric rate constants from each model can be considered directly analogous to each other if the results $S_{p,0} = \frac{d_{p,0}}{6}$ and $\varepsilon_{p,0} = 1$ for a non-porous molecule is substituted into Eq. 13 to give:

$$k_v = \frac{6k_s}{d_{p,0}\rho_M} \qquad (13)$$

It can then be observed that $k_v \approx 3(k_v)_{lm}$.

Apparatus

The P&ID of the apparatus used in the experiments is shown in Figure 2. The reactor consists of a glass tube, of ~$1.2cm$ internal diameter and $30cm$ in length, inside which the packed bed is constructed. $2g$ of a sample of soot was used in each experiment, taken from a larger batch where $8mg$ of Printex-U was diluted with $100g$ of heat-treated inert quartz sand ($150 - 200\ \mu m$ in size) to ensure consistency across all nine experiments. Diluting the soot this way reduces mass transfer limitations into the particle by breaking up large agglomerates[2], and facilitates heat transfer to ensure it remained isothermal during the exothermic oxidation reaction. The quartz sand also helps to limit the pressure drop in the packed bed and facilitates the transfer of NO_2 within the bed.

The bed is supported on quartz wool below the sample, packed sufficiently enough to prevent solid particles leaving the bed. $1g$ of pure quartz sand was placed on top of the sample, which allowed gases flowing into the reactor to be heated to the desired temperature before reacting with the soot. The resulting depth of the bed was ~$3.1cm$ and was positioned in the centre of the glass tube, heated externally by an electric furnace.

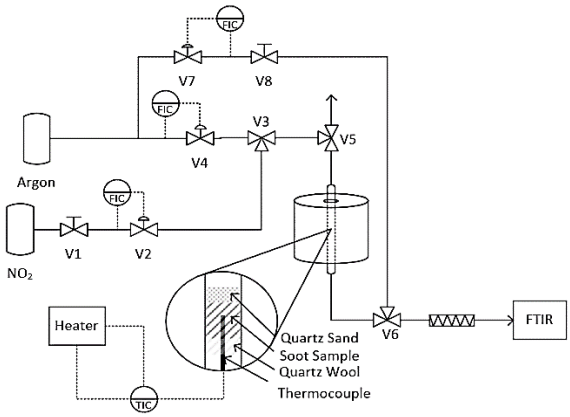


Figure 2: Experimental apparatus, flow and temperature controllers are denoted by FIC and TIC respectively.

To monitor the temperature, a thermocouple, stabilised by the quartz wool, was inserted from the base of the tube, and fed through until the tip was situated in the centre of the sample of soot, taking care to ensure it was not in contact with the edge of the tube.

A FTIR was used to determine the composition of the product stream; different molecules absorb infrared radiation at different wavelengths, which is detected by the FTIR and outputs an infrared spectrum, whereby the height of the peak is correlated with the concentration of a species, such as the Beer-Lambert law[12] where concentration is directly proportional to absorbance.

Gases exiting the reactor at ~$250\ ml\ min^{-1}$ are diluted by a factor of 8 with a flow of pure argon at ~$1750ml\ min^{-1}$, flowing through V7 and V8. The streams meet at V6 and enter the FTIR. The dilution stream increases the flowrate entering the FTIR's gas vessel and therefore decreases its residence time, τ_r. A high τ_r has the effect of delaying and smoothing the outlet concentration profile measured in the gas vessel when compared to the profile of the flow directly leaving the reactor.

Table 1: Elemental composition of Printex-U by weight percentage.[16]

C (wt%)	N (wt%)	O (wt%)
91.3 ± 0.4	0.1 ± 0.1	0.9 ± 0.2

In this study, the oxidation kinetics of Printex-U, a type of carbon black, is determined. The composition is outlined in Table 1, where it should be noted that due to its relatively low composition of organic compounds and volatiles, and a higher carbon content (at 91.3%) compared to typical diesel soots (~86.7%)[13], it is not considered to be as reactive.

Methods

Calibration of flowmeters

It was necessary to ensure that the flowmeters shown in Figure 2 were calibrated well enough to accurately reach the desired concentration of NO_2 into the reactor by altering the flowrates of argon and NO_2. These were initially calibrated on the final experimental set-up by redirecting the flow just before V5 in Figure 2 into a bubble column to determine the volumetric flowrate. Then, graphs comparing the actual flowrate against the value given by the flowmeter were plotted and the inverse of the gradient of the strongly linear relationship was implemented as a correction factor. The average error associated with the calibration of the flowrates was 4.8%.

Calibration of the FTIR

The FTIR was calibrated to able to detect the expected species exiting the reactor (i.e. $NO, NO_2, CO,$ and CO_2) with concentrations ranging from $5 - 100\ ppm$. A sample gas of known composition and concentration was passed through the system and characteristic peaks in the resulting spectrum identified using NIST[14] data. Careful consideration was taken to eliminate regions of excessive noise within the NO peak. The expected concentration of each species was plotted against the values predicted by the FTIR, and a quadratic relationship obtained for each gas. The average error associated with the calibration of FTIR concentrations was 8.48%.

Experiment preparation and heat treatment:

Prior to the start of the reaction, the system was flushed with argon for several hours to remove impurities such as atmospheric water vapour, which allowed a 'background spectrum' for the FTIR to be obtained for an 'empty' gas vessel. During the reaction, spectrums obtained are superimposed and normalised against the background spectrum to account for atmospheric contributions. This was especially important as the broad characteristic peak for water at $1630\ cm^{-1}$[14] noticeably overlaps with that of NO at $\sim 1700\ cm^{-1}$[14].

Next, the sample of soot was then heat treated by heating the packed bed to $650°C$, held at this temperature for an hour, and then cooled to the desired reaction temperature. This was necessary removed any adsorbed impurities and volatiles on the soot sample prior to the reaction (as the experiment was carried out *ex situ* and was exposed to the atmosphere before it was packed in the reactor), and to further ensure uniformity of the sample between different reactions. During this process, some CO_2 and CO was detected by the FTIR, possibly due to oxygen in the sample liberating some carbon in the sample. Since the FTIR was not calibrated for concentrations below $5ppm$, results from a previous study[2] with a similar experimental set-up was utilised to estimate the amount of carbon lost during heat treatment, amounting to ~5%. When calculating the conversion of carbon, this amount was removed from the moles of carbon initially in the bed, prior to heat treatment.

To set the desired flowrate and concentration of reactant gas, the flow through V5 was initially set to bypass the reactor. This allowed the flow to reach the desired steady state value before then being directed back through reactor, along with the spike in pressure and flow previously observed. When flow was returned to the reactor, a $5ml\ min^{-1}$ drop in argon flow was observed and accounted for when setting initial steady state values during bypass operation.

Reaction end and data processing

The reaction end was characterised as the point at which the concentrations of CO_2, CO and NO detected by the FTIR was steady at $0ppm$, at which point the system was flushed with argon to remove any remaining reactants or products in the system.

Once the concentration profile over the course of the reaction had been obtained, a finite impulse response filter (FIR) was applied to deconvolute the data and remove distortions caused by the residence time in the FTIR gas vessel. Eq. 14[15] describes the basic principle where f is the profile to be recovered, h is the original concentration profile and g is the transfer function. Using an algorithm developed previously[2], g was calculated as $[0.2\tau_r \quad 0.1\tau_r \quad 0 \quad -0.2\tau_r \quad -0.1\tau_r]$. The residence time was obtained by conducting step change experiments with the FTIR as function h, and adjusting τ_r until the deconvolved function, f, most closely resembled a step change without overshooting the steady state concentration. τ_r was calculated to be $3.09\ s$.

$$f \cdot g = h \tag{14}$$

However, deconvoluting the data amplified the noise in the concentration profile. To mitigate this, a Savitsky-Golay smoothing filter was then applied using a five-point quadratic polynomial to clarify the data.

Results and Discussion

A typical concentration profile obtained after deconvolution is shown in Figure 3. Flow of NO_2 was directed into the reactor at $t = 0s$, with a small proportion reacting, shown by the increase in concentration of the products of oxidation. The initial reaction is rapid, illustrated by the peak at $\sim 20\ s$, and then steadily decreases until $[NO_2]$ is equal to that of the inlet, and the concentration of the products fall to $0\ ppm$. The initial rate being much higher is typical for oxidation of soot reactions, typically attributed to the oxidation of the more reactive parts of the soot, e.g. adsorbed hydrocarbons.

Molar balances for carbon, nitrogen and oxygen:

Assuming a closed system and considering only the two predominant oxidation reactions, a balance for nitrogen at a time, t, can be written:

$$\dot{M}_{NO_2,in} = \dot{M}_{NO_2,t} + \dot{M}_{NO,t} \quad (15)$$

Where $\dot{M}_{NO2,in}$ is the molar flowrate of NO_2 entering the reactor and $\dot{M}_{i,t}$ is the molar flowrate of species i exiting the FTIR at time t, all in $mol\ s^{-1}$. Figure 4 indicates that no other reactions took place as the balance eventually

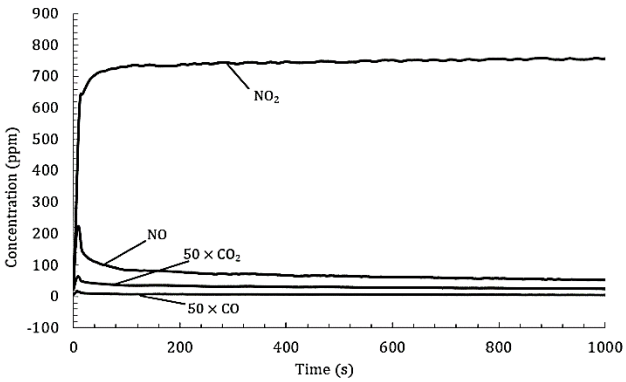


Figure 3: Deconvoluted concentration profile for oxidation reaction at $T = 400°C, [NO_2] \approx 800ppm$. For clarity, concentrations of CO_2 and CO have been multiplied by 50.

closed at $t \approx 80s$, when taking into account the average error of 2.70% with the expected moles of NO_2 entering the reactor, and 8.47% with moles of unreacted NO_2 and NO exiting the FTIR. The discrepancy in the first 80 seconds of the reaction can be attributed to the delay in NO_2 travelling through the bed and initially reacting with carbon. Assuming that the drop is purely due to reaction, it is observed that the conversion of nitrogen remains below 10% for the first ~20 seconds, indicating the reactor was operated in differential mode as the concentration of NO_2 can be assumed to stay constant.

A similar balance can be outlined for carbon, at a time, t:

$$M_{C,initial} - M_{C,t} = M_{C,oxidised,t} = \int_0^t \dot{M}_{CO_2,t} + \dot{M}_{CO,t}\ dt \quad (16)$$

Where $M_{C,initial}$ is the moles of carbon initially in the reactor bed, amounting to $\approx 1.20\ moles \pm 0.68\%$ for each experiment, and $M_{C,t}$ is the moles of carbon remaining in the bed at time t. The difference between these values represents the moles of carbon oxidised, $M_{C,oxidised,t}$, which is also equal to the moles of the products of reaction, valued by integrating the molar flowrates $\dot{M}_{CO_2,t}$ and $\dot{M}_{CO,t}$ over time.

Figure 5 demonstrates it can be assumed that the reaction progressed to completion, where all the carbon initially in the reactor exited the FTIR as CO_2 and CO for all reactions, despite the associated average error of 8.48%. This was achieved at a time analogous to the time taken for the concentration of CO_2 and CO to decrease to $\sim 0\ ppm$, demonstrated at time $t \approx 3500s$.

Finally, assuming all the oxygen in the closed system exists as NO_2, NO, CO_2 and CO, a molar balance for oxygen can be written:

$$2\dot{M}_{NO_2,in} = 2\dot{M}_{NO_2,t} + \dot{M}_{NO,t} + 2\dot{M}_{CO_2,t} + \dot{M}_{CO,t} \quad (17)$$

Similar to carbon and nitrogen, Figure 6 shows that the balance for oxygen also closed (within experimental error), indicating that oxygen atoms trapped within the particle did not oxidise the soot. It has been postulated[2] that they may have been removed initially during heat treatment, or may have been dispersed throughout the particle and released in trace amounts as CO and CO_2 over the duration of the reaction.

It is also worth noting that Figures 4, 5 and 6, demonstrate patterns repeated in all nine experiments.

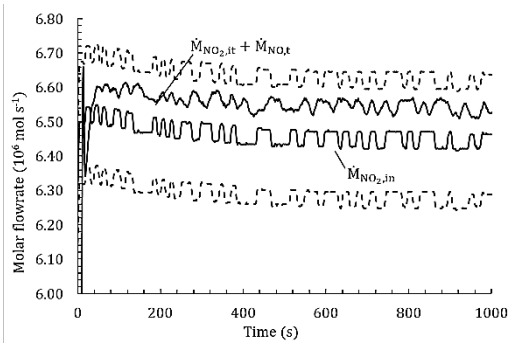


Figure 4: Plot displaying molar flowrate of NO_2 into the reactor, and unreacted NO_2 and NO over time. The dotted lines indicate the range of experimental error associated with $\dot{M}_{NO_2,in}$. This experiment was conducted at $T = 400°C, [NO_2] \approx 800ppm$.

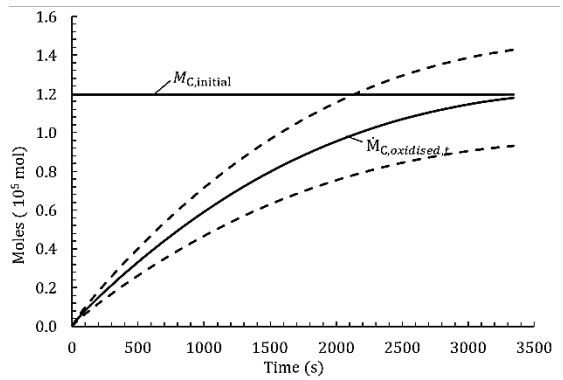


Figure 5: Plot displaying initial moles of carbon in the packed bed, and moles of carbon oxidized over time. The dotted lines indicate the range of experimental error associated with the measurements of CO_2 and CO exiting the reactor. This experiment was conducted at $T = 400°C, [NO_2] \approx 800ppm$.

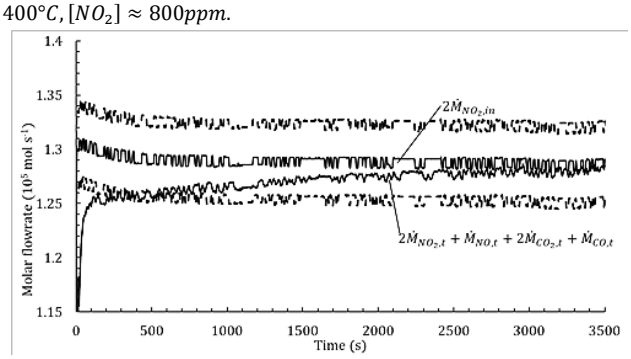


Figure 6: Plot displaying the left hand side and right hand side of the mole balance for oxygen (Eq. 17), over time. This experiment was conducted at $T = 400°C, [NO_2] \approx 800ppm$. The dotted lines indicate the range of experimental error associated with $2\dot{M}_{NO_2,in}$.

Plots of conversion against time.

Further to the molar balance, the conversion at a time, t, can be defined as:

$$X_{C,t} = \frac{M_{C,initial} - M_{C,t}}{M_{C,initial}} \quad (18)$$

Where $X_{C,t}$ is the conversion of carbon, and the numerator is equal to the moles of carbon oxidised as described in Eq. 16. The final calculated conversions ranged from 0.98 to 1.06, with an average experimental error of $\pm 11.5\%$ associated with the flowmeter and FTIR calibrations. The results were not normalised to achieve a conversion of 1 as the current results were considered as acceptable within the expected error.

Figures 7 and 8 show plots of $1 - X_{C,t}$ for different temperatures and concentrations respectively. As expected, the rate of change of conversion increases with temperature and the time to reaction completion decreases. For example, for the experiments at $800ppm$, the time taken to reach $X_{C,t} = 0.9$ increased from $t \approx 400s$ at $T = 500°C$ to $t \approx 2500s$ at $T = 400°C$. Similarly, Figure 8 demonstrates that as the concentration of NO_2 increases, the rate of change of conversion also increases. For example, for experiments at $T = 450°C$, the time taken to reach $X_{C,t} = 0.9$ increased from $t \approx 1600s$ at 496 ppm to $t \approx 800s$ at $801ppm$.

Both these trends are in accordance with trends typically reported in literature, increasing the temperature and concentration increases the likelihood of successful collisions between reactants and results in an increased rate of reaction and therefore, an increase in the rate of change of conversion. Increased temperatures also increase the likelihood of NO_2 having sufficient energy to overcome the activation energy barrier, allowing more reactions to take place.

Evaluating results against the pseudo-homogeneous model (P-S model)

Using the linearised rate equation described in Eq. 7, Figure 9 displays a plot of $\ln(1 - X_{C,t})$ against time at different temperatures. The dotted lines illustrate the construction lines for points at which the plots are no longer considered to be linear. It can be observed that the range of conversion where model is valid, generally increases with temperature. For example, at ~ 500 ppm of NO_2, the graph is linear at $400°C$ until, $X_C = 0.56$, and at $500°C$ until $X_C = 0.74$, a trend that was repeated at all three concentrations of NO_2 studied. The slope of the graph noticeably increases past these points, suggesting that the model does not represent the kinetics of the oxidation reactions well at conversions above the average across all experiments at $X_C = 0.65$ and lower temperatures.

On the other hand, summary Table 2 shows that increasing the concentration of NO_2 does not have a notable effect. For example at $T = 400°C$, the plot is linear up to $X_C = 0.56$ at a concentration of ~ 500 ppm and up to $X_C = 0.52$ and $X_C = 0.61$ for concentrations of ~ 650 ppm and ~ 800 ppm respectively. Since the average error associated with the conversions was 11.5%, the numbers were not significantly different enough to conclude that there was a clear trend.

Evaluating the results against the shrinking core model (SC model)

From the rate equation described in Eq. 12, a graph of $1 - (1 - X_C)^{1/3}$ against time was also plotted, for the same experiment in Figure 10. Similar to the pseudo-homogeneous model, increasing the temperature increased the range of conversion where the plot remained linear; for example at a NO_2 concentration of ~ 500 ppm, the plot is linear until $X_C = 0.60$ and $X_C = 0.92$ for T = 400°C and T = 500°C. Concentration did not have an observable effect on this range, except for the experiments carried out at $T = 400°C$ where X_C varied from 0.60 to 0.79 as concentration increased from $\sim 500ppm$ to $\sim 800ppm$.

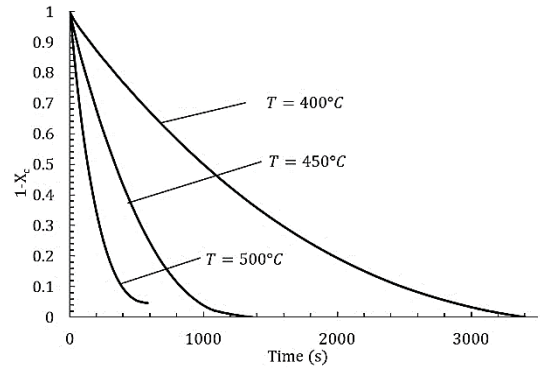


Figure 7: Plot of conversion against time for reactions at $[NO_2] \approx 800ppm$, at varying temperatures.

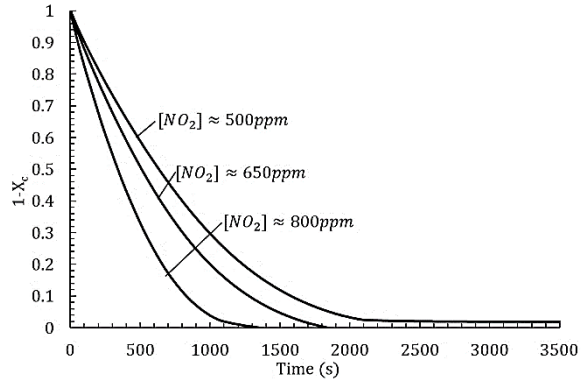


Figure 8: Plot of conversion against time for reactions at $T = 450°C$, at varying concentrations.

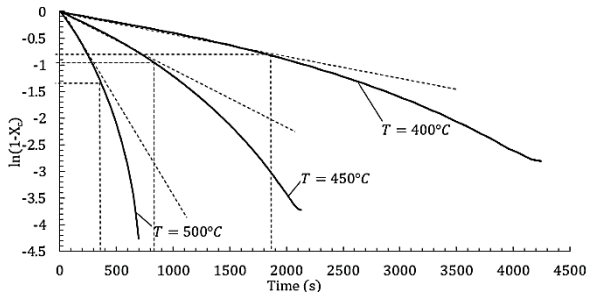


Figure 9: Plot of $\ln(1 - X_C)$ against time for various reactions at $[NO_2] \approx 500ppm$.

The average conversion where the model remains valid was calculated to be $X_C = 0.80$ for all experiments, a noticeably larger range than that of the pseudo-homogeneous model. These results suggest that the shrinking core model is a better representation of the oxidation kinetics, especially at higher temperatures and higher concentrations of NO_2.

The results in Table 2 are analogous to those obtained in a similar experiment[16] where it was concluded that at $T > 400°C$, the plot is linear for conversion up to $X_C \approx 0.5$ for the pseudo-homogeneous model, and up to $X_C \approx 0.9$ for the shrinking core model.

When the reaction begins, oxidation occurs both inside and outside the pores of the particle (comparable to reactions on its surface) which both models can account for. However, as the reaction progresses and the carbon is oxidised and the particles begin to break down, reactions may start to occur predominantly on the surface of the spherules as it has been shown that NO_2 preferentially reacts with active edge-site carbons located on the surface, which is better represented by the shrinking core model.

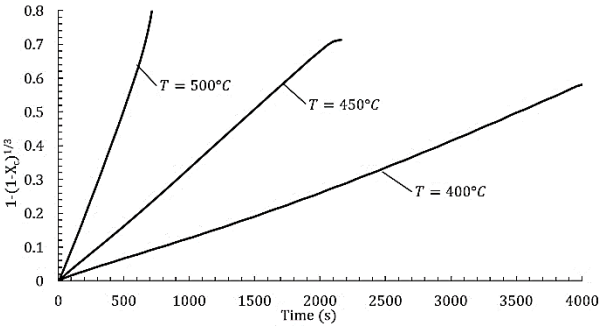


Figure 10: Plot of $1-(1-X_C)^{\frac{1}{3}}$ against time for various reactions at $[NO_2] \approx 500ppm$.

Table 2: Displaying the highest conversion at which the plot for each model remains linear.

$[NO_2]$ (ppm)	$T(°C)$	Range of X_c for P-S model	Range of X_C for SP model
500	400	0.56	0.60
	450	0.63	0.79
	500	0.74	0.92
650	400	0.52	0.70
	450	0.58	0.80
	500	0.73	0.90
800	400	0.61	0.79
	450	0.63	0.80
	500	0.84	0.88

This suggests that NO_2 is able to diffuse into the pores of the larger soot particle, but not into the spherules themselves. Although it has been suggested that the spherules within a particle may not be porous at all, an experiment[7] studying the oxidation of soot in O_2 using transmission electron microscopy, revealed that the specific external surface area of soot increased six times above the expected value from shrinkage on the external surface area of the particle alone. This was developed by another experiment[13] to determine the rate of oxidation of Printex-U in O_2, which revealed that the reaction rate reaching a second maximum after an initially fast reaction was represented well by a model describing pores inside the spherule growing and intersecting alongside the pores in the particle. This implies that unlike NO_2, O_2 was able to penetrate the spherules during oxidation.

The range of conversion for which both plots in Figures 9 and 10 remain linear increasing with temperature is a trend also observed in another study[16], which may have been down to active sites on the soot associated with a higher activation energy being more accessible to NO_2 at increased temperatures sooner. Since only three temperatures were tested, it is not possible to verify whether the average activation energy was changing with temperature, though it is noted[2] that there may have been a slight increase in overall activation energy as temperatures increased. At this point, it's possible that there are other factors making more active sites more available that are not currently considered by either model.

On the other hand, a major limitation with the current version of the shrinking core model is that Eq. 11 assumes all particles have a uniform diameter equal to $(d_p)_{lm}$, which is unlikely to be the case and different particles would 'shrink' at varying rates. Although a linear relationship was observed up to $X_C \approx 0.8$, the plots can only be considered accurate up to $X_C = 0.65$, a region shown to be unaffected by the distribution of diameters[2]. Therefore, it can be concluded that both the pseudo-homogeneous model and the shrinking core model represent the kinetics of oxidation equally well.

In addition, the assumption that the particles are spheres and assigning γ as $\frac{2}{3}$ may not be representative of particles that have been shown to have a variety of structures. In fact, it has been shown[6] that for different reactions, the value of γ varied between 0.65 and 0.8 for Printex-U.

Further to this, experimental error may have arisen at high conversions where the FTIR calibration was not sensitive to the lower concentrations of products exiting the packed bed towards the end of the reaction. The uncertainty in this region may have led to the 'wavy' regions towards the end of each plot in Figures 9 and 10.

Determining parameters k_v and n

It can be readily shown that the gradient of the plots in Figures 9 and 10, are equivalent to $-k_v C_{NO_2}^n$ and $(k_v)_{lm} C_{NO_2}^n$ respectively. The ratio of $\frac{k_v}{(k_v)_{lm}}$ was found to be an average of 3.39 across all experiments.

A graph of $\ln(k_v C_{NO_2}^n)$ against $\ln(C_{NO_2})$ can be plotted to determine the overall order of reaction, n, using Eq. 19.

$$\ln\left(k_v C_{NO_2}^n\right) = \ln(k_v) + n\ln(C_{NO_2}) \qquad (19)$$

n, was determined from the gradient as 1.13 ± 0.21. Substituting this value into an equation dividing the gradient of the plots in Figure 9 by the analogous $C_{NO_2}^n$ values allows k_v to be determined at different temperatures and concentrations.

A reaction order close to unity is typical for oxidation experiments in NO_2 above temperatures of $300°C$[6]. The ratio of $\frac{k_v}{(k_v)_{lm}}$ being higher than the

expected value of 3 could, again, be due to the distribution of the diameters of particles not being accounted for correctly, as it indicates that $(d_{p,0})_{lm}$ is higher than expected. The ratio was also consistently higher for experiments conducted at higher temperatures, which may be attributed to the particle breaking down more readily into smaller spherules, skewing the (normal[2]) distribution of diameters towards the lower end.

The Arrhenius law, outlined in Eq. 20, can then be utilised to calculate the pre-exponential factor, $k_{v,0}$ and activation energy, E_a, associated with the oxidation reactions. Here, R is equal to the molar gas constant in $J\ K^{-1}\ mol^{-1}$. A plot of ln (k_v) against $1/T$ can be seen in Figure 11, where the gradient is equal to $\frac{E_a}{R}$ and the y-intercept is equal to ln $(k_{v,0})$.

$$k_v = k_{v,0} e^{-\frac{E_a}{RT}} \tag{20}$$

$$\ln(k_v) = \ln(k_{v,0}) - \frac{E_a}{R}\frac{1}{T} \tag{21}$$

E_a was determined to be $83\ kJ\ mol^{-1}$ and $k_{v,0}$ as $112\ mol^{-1.13} m^{3.39} s^{-1}$.

The activation energy is agreeable with the range commonly reported in literature[2,16] at $50-80\ kJmol^{-1}$, but may be slightly higher due to the nature of the soot used. Experiments are typically done with diesel soot, whereas Printex-U is a type of carbon black, which has a slightly lower reactivity due to its higher proportion of graphite layer planes in the spherules, implying that there are a higher proportion of basal sites to edge sites, the former being less active and requiring higher activation energies. It is also worth noting that E_a for NO_2 is almost half than that of O_2, at $\sim 168\ kJmol^{-1}$, agreeing with values typically cited in literature.

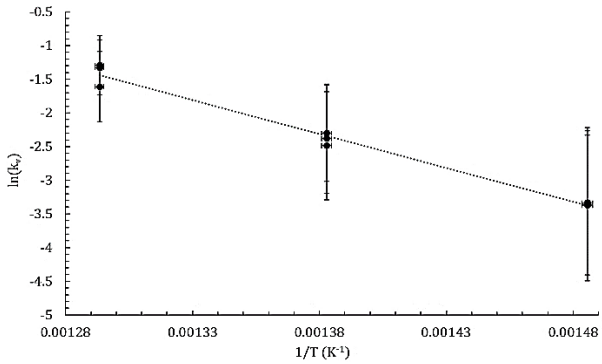


Figure 11: Arrhenius plot of ln (k_v) against $\frac{1}{T}$ using data from all experiments.

Calculating the Thiele modulus and effectiveness factor

An approximate value of the Thiele modulus for a sphere, ϕ, outlined in Eq. 22 for a first order reaction[18] can be calculated to verify the earlier assumption that the particle is not subject to diffusional limitations.

$$\phi = r_p \sqrt{\frac{k_v}{D_{e,NO_2}}} \tag{22}$$

ϕ is a function of the radius of the spherical particle (assumed), r_p, in m, the volumetric rate constant, k_v, determined earlier and D_{e,NO_2} the effective diffusion coefficient in $m^2 s^{-1}$.

Although n was determined earlier as $n = 1.13$, a first order reaction was chosen as a suitable approximation for the following calculations due to its agreement with literature data and the associated range of error. The k_v value calculated from the experiment at $\sim 800\ ppm$ of NO_2 and $T = 500°C$ was used ($= 0.27 s^{-1}$, the units are adapted for $n = 1$) as it exhibited the highest reaction rate; by attempting to remove the kinetic reaction as the rate limiting step, it was most likely to be subject to diffusional limitations. If these are not discovered to be significant, the same can be assumed for the remaining eight experiments.

The radius of the particle was approximated from the average particle size of commercial Printex-U available from The Cary Company[19] as $r_p = 1.25 \times 10^{-8}\ m$. D_{e,NO_2}[18] can be calculated as a function of the tortuosity, τ, the intraparticle void fraction, ε_p, and the pore diffusion coefficient, D_p, in $m^2 s^{-1}$.

$$D_{e,NO_2} = \frac{\varepsilon_p D_p}{\tau_p} \tag{23}$$

Tortuosity can be approximated using values calculated[20] for soot particles from various fuels. Here, the average tortuosity is reported as $\tau_p = 1.09$. Similarly, an approximation for the void fraction of the heat-treated particles was taken from previous analysis[17]. Barrett-Joyner-Halenda (BJH) analysis was employed to study the pore area and specific pore volume of a sample of Printex-U. The void fraction here is calculated as 0.43. Although the BJH method is valid only for pore diameters greater than $\sim 2nm$ (which is smaller than the diameters of the pores in Printex-U particles prior to heat treatment) due to its adoption of the Kelvin equation [9], it is unclear if pores within the spherules were also measured and neglected, therefore, some caution should be exercised with this approximation as the desired ε_p may be lower in reality.

The pore diffusion is described by the Bosanquet formula, Eq. 24[18], as a function of the molecular diffusion, D_m, and the Knudsen diffusion, D_k, both in $m^2\ s^{-1}$. The former was calculated using the Fuller method[21] as $8.46 \times 10^{-5}\ m^2\ s^{-1}$, and the latter was derived[22] using the kinetic theory of gases, as $6.88 \times 10^{-8}\ m^2 s^{-1}$.

$$\frac{1}{D_p} = \frac{1}{D_k} + \frac{1}{D_m} \tag{24}$$

In addition, the ratio of the average pore diameter and the mean free path of the NO_2, λ in m, calculated from Eq. 27[23], was much less than 0.2. Following a previous study[18], Knudsen diffusion was determined to be controlling and Eq. 24 can be reduced to $D_p = D_k$.

$$\lambda = \frac{RT}{\sqrt{2}\,\pi d^2 N_A P} \tag{27}$$

Therefore, the effective diffusion coefficient is calculated as $D_{e,NO_2} = 2.72 \times 10^{-8}\ m^2\ s^{-1}$ and the corresponding Thiele modulus as $\phi = 3.94 \times 10^{-5}$. This value can be substituted into Eq. 28 to calculate the effectiveness factor reaction for a first order reaction on a sphere.

$$\eta = \frac{3}{\phi}\left(\frac{1}{\tanh(\phi)} - \frac{1}{\phi}\right) \tag{28}$$

η was calculated as $\eta \approx 1$, indicating that the assumption that the particles are not mass transfer limited is perfectly valid.

On the other hand, due to the limited availability of similar parameters for individual spherules inside the particle, it is assumed that there are some mass transfer limitations into the pore due to their very small size[6]. It addition, it may be that O_2 is able to react within the pores of the spherule since due to its smaller kinetic diameter (0.35nm[24]) than NO_2 molecules ($\sim 0.46nm$[25]), which provides further evidence that the latter may preferentially oxidise edge-site carbons as it is less likely to diffuse into micropores on the spherule of diameter $< 2nm$, coupled with its lower bond dissociation enthalpy. Full details of calculations can be found in Appendix A.

Calculating the pressure drop in the packed bed

It is also known[18] that the diffusion regime in a particle can be affected by the pressure drop typically observed in packed bed reactors. To verify this, the Reynolds number for flow through a packed bed, described in Eq. 29[26] was used to determine the flow regime. U_s is the superficial fluid velocity calculated by dividing the volumetric flow through the bed and its cross sectional area in ms^{-1}, ρ_f is the density of the fluid, assumed to be mostly argon, in kgm^{-3}, μ is the viscosity of the fluid in $Pa\,s$, and ε is the bed voidage, approximated by comparing the volume of the quartz sand in the bed and the apparent volume of the packed bed.

$$Re = \frac{d_p U_s \rho_f}{\mu(1-\varepsilon)} \tag{29}$$

Re was determined to be 0.38, as this is less than 1, the flow is laminar[26]. Therefore, the corresponding Carmen-Kozeny[27] equation can be used to calculate the pressure drop in the bed. Here, ΔP represents the pressure drop in Pa, L is the depth of the bed in m and Φ is the sphericity of the particles in the packed bed, assumed to be 1.

$$\Delta P = -\frac{180 L\mu}{\Phi_S^2 d_p^2}\frac{(1-\varepsilon)^2}{\varepsilon^2} U_S \tag{30}$$

ΔP was calculated as $-0.18\,bar$, which was not considered to be significant enough to noticeably influence the diffusion regime into the pores of the soot particles, but further analysis may be required to determine if it is substantial enough to affect diffusion into the much smaller spherules. Micropores typically take part in configurational diffusion (due to their similar size to reacting molecules), which is considered as difficult to predict[18]. The total flowrate into the reactor was kept constant for all experiments at $\approx 250\,ml\,min^{-1}$ so it can be assumed that this value remains the same.

Determining the isothermal nature of the particle

As the oxidation reaction is rapid and exothermic, heat may not be removed rapidly enough by the fluid to keep the particle at a constant temperature (especially at the initial region where rapid oxidation occurs), so it is necessary to verify this assumption. There may be temperature effects within the particle, or within the fluid film surrounding it. It is explained[28] that for gas-solid reactions, the influence of the temperature gradient across the gas film is more likely to affect the rate of reaction. Therefore, if it can be shown that if film resistance surrounding the particle is negligible, the temperature of the particle can be considered to be the same as the temperature in the bulk and the isothermal assumption can be applied.

To calculate mass transfer resistances across the film, an iterative procedure[18] was used to calculate Δy_{NO_2}, the difference in mole fraction of NO_2 in the bulk and at the surface of the particle. Δy_{NO_2} was determined to be 8.68×10^{-17}, indicating very rapid mass transfer and no diffusional limitations across the film. Though it must be noted that the correlation $j_D(Re)$, representing the J-factor, used to obtain Δy_{NO_2} may be not be as accurate in this region of Re, but since the order of magnitude of Δy_{NO_2} is negligible, it was considered to be a suitable representation for a rough order of magnitude analysis. To verify that the temperature difference, ΔT, is also negligible, an energy balance for gases flowing in packed beds has been derived[18] in Eq. 31. It is dependent on ΔH_r the enthalpy of reaction in $J\,mol^{-1}$, c_p the specific heat capacity of the fluid (considered to be mostly argon) in $J\,kg^{-1}K^{-1}$, M_m the molecular weight of the fluid in $kg\,mol^{-1}$, $y_{f_{NO_2}}$, the film factor – a value between 0 and 1, utilising the mole fraction of reactant in the bulk and on the surface and Δy_{NO_2}.

$$\Delta T = 0.7\left(-\frac{\Delta H_r}{M_m c_p}\right)\left(\frac{\Delta y_{NO_2}}{y_{F_{NO_2}}}\right) \tag{31}$$

It can be observed that Δy_{NO_2} is directly proportional to ΔT and is sufficiently small enough to neglect temperature gradients within the particle, upholding the isothermal assumption.

Details of full calculations can be found in Appendix B.

Conclusions and outlook

The shrinking core model appears to model oxidation better than the pseudo-homogenous model in NO_2, implying that oxidation predominantly takes place on the surface of the particle. However, the shrinking core model is sensitive to the initial distribution of particle sizes. Since this is was not accounted for and a logarithmic mean approximation was used, this model is only valid for conversions under 0.65, the same range as the pseudo-homogeneous model.

It was also found that at higher temperatures both models had a better fit to the data. This could potentially be due to nitrogen dioxide attacking of a greater number of sites with higher activation energies at higher temperatures. This would limit the effect of NO_2 preferentially attacking edge sites, which may not maintain the same number of sites as oxidation occurs.

The assumptions made have also been verified through the calculation of the Thiele modulus showing that the effectiveness factor inside the particle is 1, and there is no diffusional limitations in the laminar film since the concentration of NO_2 on the surface of the particle was determined to be the same as in the bulk. Further studies

on the spherules themselves could be undertaken, since mass transfer limitations have not been evaluated. This is likely to influence rate, especially since oxygen has been shown to react homogenously throughout the soot, while NO_2 appears to oxidise only the surface of the particle.

Possible directions for future experiments include to run the experiments at a greater number and range of temperatures, this would eliminate the large errors created by uncertainty in the kinetic parameters determined and strengthen the trends observed. It may also be possible to see whether the activation energy may change with temperature, which may help justify why the models have a better fit at higher temperatures. With the increased number of data points, it may also be possible to identify a new optimal value of γ to better model the oxidation.

It is also necessary to verify whether the shrinking core model is better than the pseudo-homogenous by calculating the particle size distribution of Printex-U. The shrinking core model is sensitive to the particle size distribution, and accounting for it when calculating kinetic parameters would allow for the model to be validated across a larger range of conversions.

Mass transfer limitations can also be tested experimentally by running experiments with increased flowrates. If the rate of conversion and kinetic parameters are identical, this would provide further evidence that there are no mass transfer limitations in the bed itself. It is also possible to gain a better approximation for the mass of carbon loss before reaction by heat treating a large mass of Printex-U, and then measuring the moles of CO and CO_2 produced over time, to increase confidence in the conversion of carbon.

References

[1] Leistner, K. (2012). Experimental and Modelling Study of Catalytic Diesel Soot Oxidation. [online] Available at: tinyurl.com/ryzj3zw

[2] Tighe, C., Twigg, M., Hayhurst, A. and Dennis, J. (2012). The kinetics of oxidation of Diesel soots by NO2. Combustion and Flame, [online] 159(1), pp.77-90. Available at: tinyurl.com/yxxyhv5v

[3] Neeft, J., Smakman, E., Makkee, M., Moulijn, J. and Nijhuis, T. (1997). Kinetics of the oxidation of diesel soot. Fuel, [online] 76(12), pp.1129-1136. Available at: tinyurl.com/ug3l3cy

[4] Stanmore, B. (1991). Use of the differential thermal gravimetric analyser for carbon/oxygen reactivity measurement. Fuel, [online] 70(12), pp.1485-1487. Available at: tinyurl.com/t42bykj

[5] Gilot, P., Bonnefoy, F., Marcuccilli, F. and Prado, G. (1993). Determination of kinetic data for soot oxidation. Modeling of competition between oxygen diffusion and reaction during thermogravimetric analysis. Combustion and Flame, [online] 95(1-2), pp.87-100. Available at: tinyurl.com/ruwomzt

[6] Stanmore, B., Brilhac, J. and Gilot, P. (2001). The oxidation of soot: a review of experiments, mechanisms and models. Carbon, [online] 39(15), pp.2247-2268. Available at: tinyurl.com/u6urnmc

[7] Toth, P., Jacobsson, D., Ek, M. and Wiinikka, H. (2019). Real-time, in situ, atomic scale observation of soot oxidation. Carbon, [online] 145, pp.149-160. Available at: tinyurl.com/wawwl3k

[8] Singh, M., Srilomsak, M., Wang, Y., Hanamura, K. and Vander Wal, R. (2018). Nanostructure changes in diesel soot during NO2–O2 oxidation under diesel particulate filter-like conditions toward filter regeneration. International Journal of Engine Research, [online] 20(8-9), pp.953-966. Available at: tinyurl.com/yyxlu6bw.

[9] Müller, J., Frank, B., Jentoft, R., Schlögl, R. and Su, D. (2012). The oxidation of soot particulate in the presence of NO2. Catalysis Today, [online] 191(1), pp.106-111. Available at: tinyurl.com/upant7q.

[10] Ishida, M., Wen, C. and Shirai, T. (1971). Comparison of zone-reaction model and unreacted-core shrinking model in solid—gas reactions—II non-isothermal analysis. Chemical Engineering Science, [online] 26(7), pp.1043-1048. Available at: tinyurl.com/vo93s9e.

[11] Salgado, M. and Rossi, M. (2002). Flame soot generated under controlled combustion conditions: Heterogeneous reaction of NO2 on hexane soot. International Journal of Chemical Kinetics, [online] 34(11), pp.620-631. Available at: tinyurl.com/tu9jd5d.

[12] Mathias, J. (2019). How Does FTIR Work? | Innovatech Labs. [online] Innovatech Labs. Available at: tinyurl.com/see9xxl

[13] Tighe, C., Twigg, M., Hayhurst, A. and Dennis, J. (2016). The kinetics of oxidation of Diesel soots and a carbon black (Printex U) by O2 with reference to changes in both size and internal structure of the spherules during burnout. Carbon, [online] 107, pp.20-35. Available at: tinyurl.com/vqo64ah.

[14] Webbook.nist.gov. (2019). NIST WebBook. [online] Available at: tinyurl.com/y8xduyye

[15] En.wikipedia.org. (2019). Deconvolution. [online] Available at: tinyurl.com/y7kederw

[16] Chen, Y. (2019). Investigation of Microstructure Modification in Carbonaceous Nanoparticles During Low Temperature Oxidation by NO2.

[17] Hu, T. (2017). Experimental Investigation of Pore Growth in Carbonaceous Nanoparticles during Low Temperature Combustion.

[18] Froment, G., De Wilde, J. and Bischoff, K. (1990). Chemical reactor analysis and design. Hoboken, NJ: John Wiley & Sons.

[19] The Cary Company. (2019). PRINTEX-U Powder - RCG, Carbon Black for Inks, Polymers & Coatings. [online] Available at: tinyurl.com/uhu69qw

[20] Savic, N., Rahman, M., Miljevic, B., Saathoff, H., Naumann, K., Leisner, T., Riches, J., Gupta, B., Motta, N. and Ristovski, Z. (2016). Influence of biodiesel fuel composition on the morphology and microstructure of particles emitted from diesel engines. Carbon, [online] 104, pp.179-189. Available at: tinyurl.com/svuowra.

[21] Umich.edu. (2019). Prediction of Binary Gas Diffusivities. [online] Available at: tinyurl.com/r4d56s3

[22] Welty, James R.; Wicks, Charles E.; Wilson, Robert E.; Rorrer, Gregory L. (2008). Fundamentals of Momentum, Heat and Mass Transfer (5th ed.). Hoboken: John Wiley and Sons.

[23] Physics.bu.edu. (2019). Molecular Collisions. [online] Available at: tinyurl.com/yx7ngdh9

[24] Kentish, S., Scholes, C. and Stevens, G. (2010). Carbon Dioxide Separation through Polymeric Membrane Systems for Flue Gas Applications. Recent Patents on Chemical Engineering, [online] 1(1), pp.52-66. Available at: /tinyurl.com/sxh4al6.

[25] Watson, A. and Valberg, P. (2001). Carbon Black and Soot: Two Different Substances. AIHAJ - American Industrial Hygiene Association, [online] 62(2), pp.218-228. Available at: tinyurl.com/u5m2vn5.

[26] Neutrium.net. (2019). Packed Bed Reynolds Number – Neutrium. [online] Available at: tinyurl.com/vagprzn

[27] McCabe, Warren L.; Smith, Julian C.; Harriot, Peter (2005), Unit Operations of Chemical Engineering (seventh ed.), New York: McGraw-Hill, pp. 163–165

[28] Levenspiel, O. (1998). Chemical reaction engineering. Wiley.

A study into the efficacy and durability of eggshell based superhydrophobic coatings formulated under different conditions

Whai Shin Ooi and Pichan Tangtrakulwong
Department of Chemical Engineering, Imperial College London, U.K.

Abstract A facile fabrication of an eggshell-based superhydrophobic coating was successfully performed and analysed in this work. The calcite-based eggshells were functionalised through the adsorption of stearic acid, and through mixing with a polyethylene/toluene binder solution, a paint suspension was formed. Different application methods (spin coating, dip-coating and pipette coating) were applied to various substrates (glass, copper, aluminium, sandpaper and filter paper) with and without slide pre-treatment and eggshell deproteinisation, and the effect on superhydrophobicity was quantified through the measurement of the contact angle. All coated substrates showed superhydrophobicity (with a contact angle greater than 150°) except filter paper, but for pipette coating only. Substrate pre-treatment was found to have little effect on superhydrophobicity, but eggshell deproteinisation had decreased hydrophobicity. Chemical and mechanical durability tests were conducted by immersion in different pH solutions for extended periods of time and abrasion with a magnetic stirrer respectively. Contact angles decreased over time during chemical stability tests, with a slightly greater decrease for solutions with a pH further from 7. Mechanical durability was found to be substrate-dependent, with an almost total loss of coating from aluminium and copper, partial loss from sandpaper, and no significant decrease in contact angle or coating on the glass substrate.

Introduction

Hydrophobicity is a property of a surface characterised by the contact angle with a water droplet at rest: a lower contact angle is associated with greater hydrophilicity and a higher contact angle is associated with greater hydrophobicity (Figure 1). Superhydrophobicity is defined through the following properties: a high contact angle with water (>150°), low contact angle hysteresis (<10°), and a low roll-off angle (<5°). This property arises from the combination of low surface energies and hierarchical structures on the surface, resulting in high surface roughness. The Cassie-Baxter and Wenzel models propose that the microscale structure from the asperities in the surface topology, as well as the nanoscale structure due to the functionalisation of the surface, both contribute to the surface roughness and hence superhydrophobicity f(Simpson, Hunter & Aytug, 2015). The asperities filled with air prevent complete wetting of the coating surface, thereby reducing all interactions with the surface, whereas the presence of the hydrophobic ends of the fatty acid chains exhibits unfavourable interactions with water. A lower water surface energy is therefore achieved when there is a lower contact area between the water droplet and surface, equivalent to a higher contact angle (Figure 2).

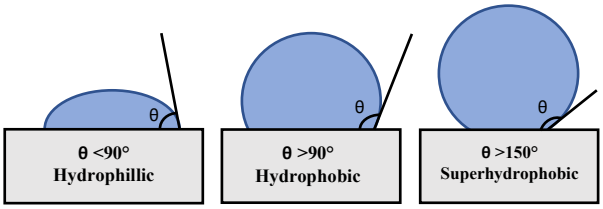


Figure 1 Wettability and contact angles of solid surfaces

There have been efforts in the last decade to develop superhydrophobic coatings which, when applied to a surface could demonstrate the ability to repel any water that encounters it. This property is very useful and has commercial applications in making surfaces self-cleaning, corrosion-resistant or anti-bacterial. Increased attention has been paid to superhydrophobic surfaces that exhibit multifunctional properties, combining their water-repelling characteristics with additional mechanical, optical or electrical functions (Jeevahan et al., 2018). Although superhydrophobic paints have been fabricated using various techniques and raw materials, many obstacles prevent its commercial success, namely concerns regarding the cost of the fabrication process, environmental concerns and the durability of the final coated product. More recently, research has been geared toward more facile and inexpensive methods of developing superhydrophobic paints, using more environmentally friendly raw materials, such as eggshell wastes.

Calcium carbonate powder is very cheap and widely used as surface coatings in the industry. Many successful studies have been able to use it as a raw material to fabricate superhydrophobic coatings. However, the increasing emphasis on environmental sustainability makes it imperative to use an alternative source. About 250,000 tons of eggshell waste is produced annually (Verma, Kumar & Bansal, 2012), and has a calcium carbonate content of 96-97% (Cree & Rutter, 2015). This makes it an ideal substitute to commercial calcium carbonate, typically produced through mining and refining, reducing the costs of production and reducing environmental damage. Research has been done to demonstrate its viability: Wen, Huang & Guo (2019) produced superhydrophobic coatings using stearic acid-modified eggshell powder, zinc oxide, ethanol and carboxymethyl cellulose as an adhesion agent. Causby (2019) compared the effects of

different fatty acids on the superhydrophobicity of the paints and optimised the ratios between the eggshell and the fatty acid as well as the functional powder to binder solution. Fatty acids, such as stearic acid, are commonly used as a surfactant for phase and surface modification of calcium carbonate particles. Stearic acid reacts with calcium carbonate to form adsorbed calcium stearate which achieves hydrophobicity due to its long fatty acid chain and lowering of the calcium carbonate surface energy (Shi, Rosa & Lazzeri, 2010). Moreover, functionalisation of calcium carbonate particles with fatty acid alters its morphology by selectively forming the aragonite phase over calcite phase which increases the degree of roughness leading to a higher contact angle. (Sarkar, Ghosh & Mahapatra, 2012).

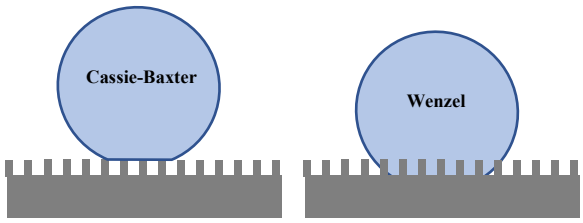


Figure 2 Wetting states, Cassie-Baxter and Wenzel models

While much of research on superhydrophobic paint fabrication has been carried on glass slides, it is important to note that the potential industrial applications of these superhydrophobic paints span beyond this substrate. A notable substrate that would benefit from the application of the paint is on metal surfaces, especially copper and aluminium, that often undergo corrosion and tend to accumulate ice in cold weathers. Having a durable superhydrophobic coating over the metal surfaces could result in increasing the lifespan of apparatuses made from these substrates. There has been significantly less research into the application of superhydrophobic paints on porous and hydrophilic substrates, such as paper and fabric, which are penetrable by water. The ability to induce superhydrophobicity on such substrates could open opportunities for many applications. For example, superhydrophobic cellulose paper was fabricated which can be used in the paper packaging industry, as wettability is important as paper and paperboard lose its rigidity as a result of absorbing water during distribution and storage (Wang et al., 2017). Filter paper coated with superhydrophobic coating was also able to demonstrate excellent oil/water separation efficiency (He et al., 2018).

There have been numerous methods developed to coat superhydrophobic paints onto substrate surfaces, with varying difficulty and effectiveness. Zhang et al. (2011) fabricated a superhydrophobic surface by spin coating glass with a suspension of calcium carbonate modified with stearic acid which demonstrated self-cleaning properties. Another facile method that is commonly used in research is the sol-gel dip-coating method (Causby, 2019). Other more complicated application methods include electrodeposition, photolithography and spray-drying, but they do not demonstrate the same benefits of simplicity that the first two methods possess. This report attempts to compare the effectiveness of three facile methods of paint coating in inducing superhydrophobicity: dip coating, spin coating, and pipette coating.

For effective coating, the paint must be able to form strong adhesive forces with the substrate surface and sufficiently cover the surface. Impurities on the substrate surface can interfere with the interactions between the paint coating and surface, leading to weaker adhesion. An example is a hydrocarbon layer on the surface can lead to a low nucleation density of the coating film and a lower adhesion of the coating to the surface. Additionally, particulate contaminants can also lead to pinholes forming in the coating (Martin, 2015). Therefore, it is crucial to ensure that the substrate is clean and smooth, without any contaminants. Acetone and ethanol are standard organic solvents which are used to clean glassware, not only are they very polar substances which dissolves almost all organic compounds, but they are also miscible in water and can be easily washed off. Pre-treatments of eggshells have also been carried out in previous works to enhance the purity of calcium carbonate in the eggshells (Causby, 2019), as impurities may lead to undesirable side reactions with the fatty acids or binder solution. A similar methodology was used to deproteinise shrimp shells for the extraction of chitosan performed by Queiroz Antonino et al. (2017) and Younes & Rinaudo (2015). This work studies the effect of the pre-treatments on substrates and eggshells prior to coating fabrication on its superhydrophobicity.

It has been repeatedly demonstrated that glass slides coated with superhydrophobic paints have excellent UV, chemical and mechanical durability, displaying self-cleaning and anti-icing properties (Wen, Huang & Guo, 2019; He et al., 2018). However, one of the greatest obstacles to the commercial development of these superhydrophobic paints is its chemical and mechanical durability. Little attention has been paid to this area of research until recently, and even less so regarding the ability for the paints to withstand chemical and mechanical stresses when applied to different substrates. He et al. (2018) fabricated superhydrophobic material using eggshell powder, stearic acid, zinc oxide and ethanol and demonstrated that the paint could maintain superhydrophobicity on glass slides after mechanical and pH durability tests. By attempting to understand how the superhydrophobicity is affected and the mechanisms that contribute to wearing of the paint's properties, improvements can then be made to increase its durability under these conditions.

This study focuses on demonstrating the universality of Causby's (2019) optimised formulation of superhydrophobic coating from eggshell waste on substrates of different properties: non-porous and smooth (glass, copper, aluminium), non-porous and rough (sandpaper), as well as porous (filter paper). The resistance of the paints on the different substrates to

varying pH conditions and mechanical abrasion over time was also investigated. A comparative study on the different facile methods of paint application was carried out to determine the simplest and effective method on the different substrates as well. The work also attempts to look into the effect of pre-treatment attempts of the glass slide and eggshells on hydrophobicity. These results will serve as the platform for discussion about the interactions between the coating surface and water, as well as that between the coating surface and substrate. Through doing so, a careful analysis of the mechanisms that lead to the formation and degradation of the superhydrophobicity of eggshell paints can be carried out in this paper.

Methodology

Eggshell Powder Preparation

Chicken eggshell waste was obtained from a local Thai restaurant (Kobkhun Thai Cuisine). The shells were washed with cold water for 15 minutes to remove any remaining organic matter, before being boiled at 100 °C for 1 hour to remove the inner membrane and any other impurities. Remaining membranal tissue was then peeled off the eggshells and left to dry at room temperature overnight. After drying, the eggshells were crushed into a fine powder using a commercial blendir (AmazonBasics) and sieved in an electromagnetic sieve shaker (Endecotts) for 20 minutes to obtain an eggshell powder of particle size smaller than 32 µm. This is in accordance to Causby's (2019) paper where this range of particle size displayed the highest contact angle when treated with stearic acid.

Pre-treatment Methods

The ground eggshell powder was mixed with 1 M sodium hydroxide (BPH Chemicals) on a magnetic stirring plate (Cole-Parmer) for 1 hour to remove and denature any remaining proteins from the eggshell membranes. Following deproteinisation, the powder was filtered and washed with DI water in cycles until a pH 7 was reached in the suspension. This was measured using pH indicator paper (Advantec). The powder is then dried overnight before used in paint formulation.

The effect of substrate pre-treatment was carried out by subjecting the substrates to different conditions: (1) in acetone for 15 minutes, (2) in acetone then ethanol, for 15-minute periods each, (3) coating with a layer of nail polish and left to dry. After pre-treatment, the slides were washed thoroughly with DI water before being dried in the laboratory oven (Memmert) at 90 °C and a vacuum pressure of 350 mb for 1 hour.

Paint Formulation

To induce superhydrophobicity, the eggshell powder was grounded for 15 minutes with 2 wt% stearic acid granules (BPH Chemicals) using a mortar and pestle. This is based on previous work which demonstrated this ratio of eggshell to fatty acid to be optimal for functionalisation (Causby, 2019; Jafari & Farzaneh, 2012). The binder solution was prepared using a stirring hot plate setup (Camlab), in a 500 mL round-bottomed flask under reflux, by dissolving polyethylene grains (BPH Chemicals) into toluene (BPH Chemicals) at a ratio of 2 w/v% (PE/toluene) to form the optimum adhesive solution for eggshell paint formation (Kokkanis, 2017). This was completed at a temperature of 90 °C and a rotational speed of 600 RPM for 2 hours.

The functional powder was then added to the binder solution at a solid content of 40 w/v% (powder/binder), also based on Causby's (2019) findings. Mixing was carried out in a sealed flask using a magnetic stirrer (Cole-Parmer) at 80 °C and 400 RPM for another 30 minutes for further functionalisation of the eggshells and dissolving of polyethylene. Finally, the mixture was cooled to room temperature, while under magnetic stirring, to produce the paint suspension.

Substrate Coating

Substrate samples (glass, aluminium, copper, sandpaper and filter paper) were cut into identically sized pieces (1.5 x 2.5 cm) and attached to a glass slide using double-sided tape to ensure a flat surface. Three different methods were used to coat the substrates with the superhydrophobic paint: pipette coating, dip coating and spin coating. For pipette coating, 300 µL of paint was pipetted (Cole-Parmer) onto the substrate surfaces, resulting in a paint thickness of 0.8 mm. For dip coating, the substrates were inserted into a 50 mL beaker of the paint suspension and tilted at an angle of 30° for 15 seconds before removing it. The coating was also applied via the spin-coating method: 300 µL of paint was pipetted onto the substrates, placed onto the platform of the spin coater (Laurell Technologies) and spun at 750 RPM for 2 minutes. All coated samples were left overnight to dry on a horizontal surface before contact angle measurements were made.

Contact Angle Measurement

The substrate coatings were characterised through contact angles measured digitally using an automated goniometer setup (Rame-Hart) as shown in Figure 3. The coated samples are placed onto the projection platform and a high definition camera is used to reflect the lateral view of the surface to set the baseline on its complementary software (DROPimage Advanced). A 10 µL drop of deionised water was then placed onto the surface, and the camera captured the side-profile of the droplet, taking 50 consecutive measurements of the left and right contact angles. This is repeated at 5 different locations to find an average angle reading.

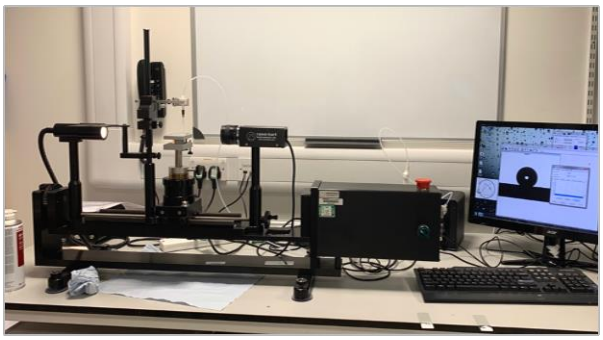


Figure 3 Goniometer setup

Chemical Stability Test

The pH solutions were prepared by adding either hydrochloric acid (BPH Chemicals) or sodium hydroxide (VWR chemicals) into deionised water. The exact pH was measured by a pH meter (Mettle-Toledo) to produce solutions of pH values 3, 5, 7, 9 and 11. Chemical stability tests were carried out by immersing the coated substrate samples in 20 mL of different pH solutions for 15-minute periods, followed by drying in an oven (Memmert) at 60 °C and vacuum pressure of 350 mb for 45 minutes. The contact angles were then measured after 0, 15, 30, 45, 60 minutes of submergence into the solutions and repeated for all substrates.

Mechanical Durability Testing

To perform mechanical durability tests on the coated substrates, the samples were placed into a beaker with a magnetic stirring bead resting on the coated surface. The beaker was then placed on top of a magnetic stirrer plate (Stuart) and operated at 500 RPM for 30 minutes, which caused the bead to rotate on the substrate surfaces. After the abrasion tests have been completed, any remaining coating flakes or dust impurities were removed. The contact angles were measured before and after the durability tests.

Results and Discussion

This section conveys all observed results gathered through variations in fabrication methods of the superhydrophobic eggshell coatings as well as its durability tests. The implication of these results in terms of potential physical meanings and impacts were discussed in detail. It is important to note that the apparent static contact angle is referred to as 'contact angle' for simplicity.

Comparison of using precipitated calcium carbonate and eggshell powder as filler

The superhydrophobicity of eggshell based coatings compared to ones produced from precipitated calcium carbonate was compared. Causby (2019) compared the X-ray diffraction patterns between precipitated calcium carbonate and eggshell based coatings to demonstrate the high concentration of calcium carbonate in eggshells, and thus its viability as a substitute. This was verified through the superhydrophobic properties of both paints in this study. Coating fabricated using functionalised powder with a size range smaller than 32µm was determined to exhibit the highest contact angle, as it leads to a more uniform exposure of the hydrophobic tails of stearic. Although larger eggshell particles were expected to provide a rougher microscale topology leading to a higher contact angle, the surface coverage with fatty acid was found to be dominating.

As shown in Figure 4, both untreated eggshell and precipitated calcium carbonate based coatings displayed superhydrophobicity, with a water contact angle greater than 150°. The results show that eggshell was just as effective as its pure counterpart when made into a superhydrophobic coating. It could also be implied that any presence of impurities in the eggshell did not significantly affect the hydrophobicity of the coating. An important observation made was that large standard deviations in the contact angle measurements were found. Vuckovac et al. (2019) reported that this can be attributed to large uncertainty in the precision of the goniometer. A misplacement of the baseline by as little as 1 µm can lead to an error of up to 10° in the contact angle measurement. This made measurements of the contact angles of water droplets on substrates that have visible roughness on its surface very difficult and prone to error.

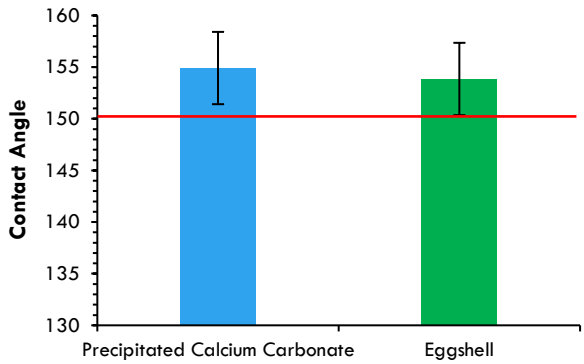


Figure 4 WCAs of coatings made by precipitated calcium carbonate and

Comparison of coating application methods on different substrates

From the different methods of coating investigated, pipette coating was the most effective in producing the best results in terms of superhydrophobicity. This method was therefore used for all coatings in other experiments. There is potential to optimise the uniformity of the paint surface through better understanding the mechanisms of the coating processes. This requires further characterisation of the paint, especially in terms of the physical properties of the bulk liquid.

As seen in Figure 5, coating with pipette method produces the highest contact angle on glass, copper,

aluminium and sandpaper, and was the only method that was able to fabricate a coating that exhibited superhydrophobicity (contact angle > 150º). It was observed that the only exception to this was filter paper. As filter paper is a porous substrate consisting of microfibers, the applied coating seeps into the pores resulting in a thinner and more uneven coating. One observation is that the paint diffused into the filter paper before being able to spread and wet the whole substrate surface area. The same result was obtained for all coating methods, demonstrating limitations of the application of the eggshell based paint in this study on porous substrates. Assuming the coating was applied evenly so that a macroscopically smooth surface was produced, the pipette coated paint has a thickness of 0.8 mm. This is significantly higher than that of the coatings produced through the other methods: spin and dip coating. As the paint is merely pipetted onto the surface of the substrates, the thickness of the film is determined dominantly through the ability of the paint liquid to wet the surface. The property is affected by both the surface energies of both the liquid and substrate surface: the fact that the paint has spread to cover all the substrate surfaces shows that surface tension of the eggshell paint is lower than the surface energies of the substrates.

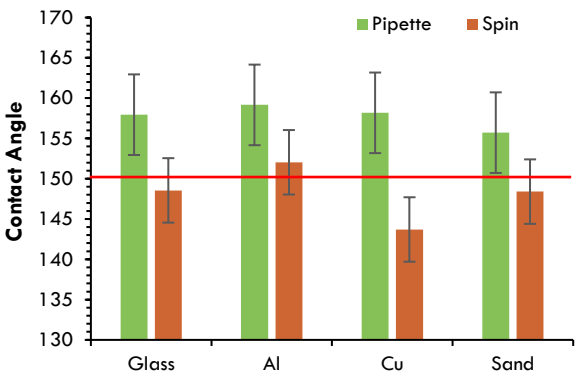


Figure 5 WCAs of different substrates coated by pipette coating and spin coating

Another important factor that contributes to the formation of the paint coating is the drying process. The evaporation of toluene from the paint suspension at the surface and the concurrent coalescence of the calcium carbonate, calcium stearate, polyethylene mixture has the ability influence the surface tension, viscosity and density of the paint. The drying mechanism causes a concentration gradient throughout the film and ultimately affects the surface properties of the paint (Alsoy & Duda, 1998), namely its superhydrophobicity in the context of this work. The difference in concentrations can also be used to explain the decrease in contact angle measured after mechanical abrasion of the paint surface. Techniques such as ATR-FTIR spectroscopic imaging can be used to further study the drying mechanism of the paint and its surface homogeneity.

Spin coating of substrates was conducted as a comparison to pipette coating, to investigate its effectiveness in fabricating a superhydrophobic paint. The spin coating makes use of centrifugal forces (Zhang & Hoshino, 2019) to produce coatings of thicknesses in the range of 1 to 10 µm (Smith, Inomata & Peters, 2013). Although the volume used for spin coating was fixed to be identical to that used in the dip coating method, it was clear from observations that the method was unable to produce a uniform coating on the substrates. As spin coating only uses 2-5 % of the coating deposited on the substrate while the remaining 95-98 % was flung off and disposed of (Tyona, 2013), this method of the coating was not as efficient in terms of the volume of coating utilised.

There was clear evidence of Wafer edge effects on spin coated substrates as observed in Figure 6b, where a high concentration of particles was seen to have accumulated at the edges of the substrates, as well as the middle where the liquid was deposited. This is largely attributed to the surface tension effects of the liquid paint as it is spun on the platform, preventing the suspension from spreading out on the substrate surface. A rim of thicker paint coating was formed around the perimeter of the substrate, resulting in non-uniformity in film thickness on the surface. The square geometry of the substrate itself disturbs the centrifugal air flow during its spin and contributes to the non-uniformity of the paint surface, especially on the corners of the substrate (Tyona, 2013). It is also noted that the functional powder (calcium carbonate, calcium stearate and stearic acid) has a density which is three times larger than the binder solution (toluene and polyethylene) (National Center for Biotechnology Information, 2004). Therefore, this forces the denser functional powder particles towards the rim of the substrates through centrifugation.

Thus, there is potential for further optimisation of the spin coating method with respect to its spin velocity to promote better uniformity around the edges of the substrates. This means that it is crucial to understand the dependence of spin speed and viscosity on the coating thickness and its relationship to the rate of evaporation and film formation (Tyona, 2013). Moreover, the use of a circular substrate could reduce the wafer edge effect exhibited, leading to a more uniformly coated surface.

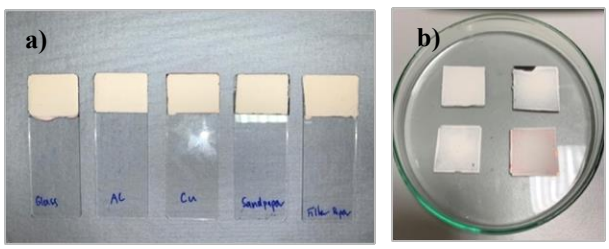


Figure 6 Substrates coated by a) pipette b) spin coating

Application of the eggshell paint was attempted using the dip coating method as well, although this

method gave very thin and uneven coatings on the substrates. The critical step of the dip coating method is the withdrawal process, where the substrate is removed from the paint suspension and a thin layer of paint remains on its surface. Formation of the paint coat and its subsequent thickness are dependent on the balance between two main forces: a draining force that draws the suspension back into the bulk liquid due to gravity, and an entraining force that adheres and retains the fluid on the surface (Ossila, n.d.). It can be inferred from the observation that the low viscosity of the liquid paint has led to a dominating entraining force to form a thin and uneven coating.

The unevenness of the surface was visible through the high definition camera, which made the determination of the baseline and thus measuring contact angles extremely difficult for samples namely sandpaper and filter paper. The other coated substrates were unable to demonstrate superhydrophobic properties; SEM analysis can be carried out on the coating formed to characterise the surface and allow for a better understanding of the phenomenon. It was also observed that dip coating produced samples with low mechanical durability that can be easily rubbed off by hand. For future works, the physical properties of the wet paint, namely viscosity and density, should be measured to gain a better understanding of the coating regime and drying mechanism, and thus better quantify the thickness of the film formed.

From the observations correlating the thickness of the paint coating and its superhydrophobic property, it is worth further investigations into the existence of a minimal thickness of paint that produces superhydrophobicity, but thick enough to withstand chemical and mechanical stresses. This requires an in-depth investigation and characterisation of how the paint film is formed and how superhydrophobicity is affected. By understanding the critical parameters that can be used to control the process, then it can be engineered to fit industrial applications, although this is beyond the scope of this work.

Chemical Stability Testing

After submerging the paint samples in different pH solutions for 1 hour, all the substrates displayed similar general trends (Figure 7). The chemical stability of coated filter paper was not tested as it was difficult to fabricate a coated surface with the desired superhydrophobic properties. While contact angles were observed to show a decreasing trend over time following immersions in all pH solutions, immersions in pH 7 solution led to the smallest decrease in contact angles. It has been previously determined that the decrease in contact angle in neutral, acidic and basic solutions were due the change in surface chemistry and morphology (Seeharaj et al., 2019), which can be used to explain the trends observed.

It has been found that stearic acid deposited on the coated surface is partly chemisorbed (forming calcium stearate) and partly physisorbed (Shi, Rosa & Lazzeri, 2010). Kokkanis (2017) also showed using thermogravimetric analysis of the functional powder that there was unreacted stearic acid which was physisorbed on the surface. In basic solutions, sodium hydroxide is likely to neutralise the physisorbed stearic acid on the substrate surface producing sodium stearate and water. Sodium stearate is highly soluble in water (American Elements, n.d.), and thus its dissolution will lead to a decrease in the hydrophobicity of the coating. In acidic solutions, protons from hydrochloric acid will react with calcium stearate or calcium carbonate, producing calcium chloride and stearic acid, or calcium chloride and carbon dioxide respectively. Both mechanisms contribute to the reduction of the effective active area of calcium carbonate for functionalisation by stearic acid. These would result in the destruction of the hierarchical structure to a smoother surface, thereby also contributing to a decrease in hydrophobicity.

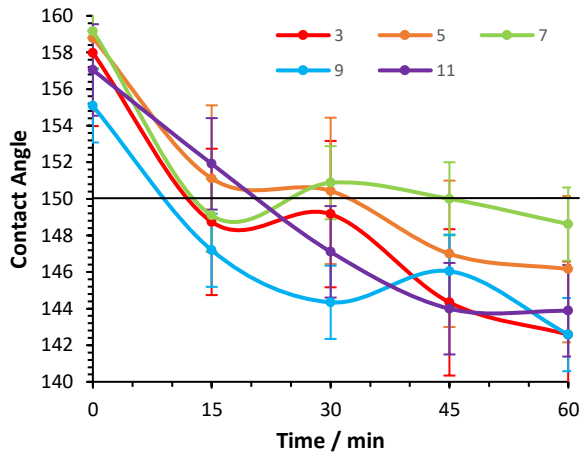


Figure 7 WCAs of coated aluminium in different pH environments after immersions over time

The decrease in superhydrophobicity when submerged in deionised water (pH 7) suggests that there were contributions other than reactions from the acidic/basic environments that could affect the property. Although stearic acid has an aqueous solubility of 1.08 mg/100 mL (National Center for Biotechnology Information, 2004), there was considerably less than 2.4 mg of stearic acid (total in one paint sample) exposed on the surface of the paint, and so a significant proportion of it is likely to have dissolved in the solutions during immersion. Chemical characterisation techniques such as ion chromatography can be used to determine the chemical compositions of the pH solutions following chemical stability tests to better understand the reaction mechanisms that occur during the submergence of paint samples.

Another reason for an observed decrease in contact angles could be due to experimental limitations. It is difficult to ensure that moisture was completely removed when the substrates were oven-dried after each immersion. This is because of the hierarchical roughness of the paint surface, in which moisture could be trapped.

The presence of moisture in the grooves of the coating will lead to a higher adhesion towards the water droplet during contact angle measurements compared to a completely dry surface. The favourable interactions between the water in the grooves and the water droplet from the goniometer increase wetting of the overall surface, leading to a lower contact angle.

It is interesting to note that despite the effect of several proposed mechanisms on the decrease of superhydrophobicity, the difference in the final contact angle for different pHs is not apparent, as most final contact angles after 60 minutes fall within the standard deviation of one another. Good surface coverage of the eggshell particles by chemisorbed stearic acid can be inferred, as illustrated in Figure 8b. The nanoscale schematic demonstrates that the densely packed, long hydrophobic fatty acid chains that covered the substrate surface provide high steric hindrance, making it more difficult for protons and hydroxide ions to attack the calcium carbonate or the functional end of the stearic acid molecules.

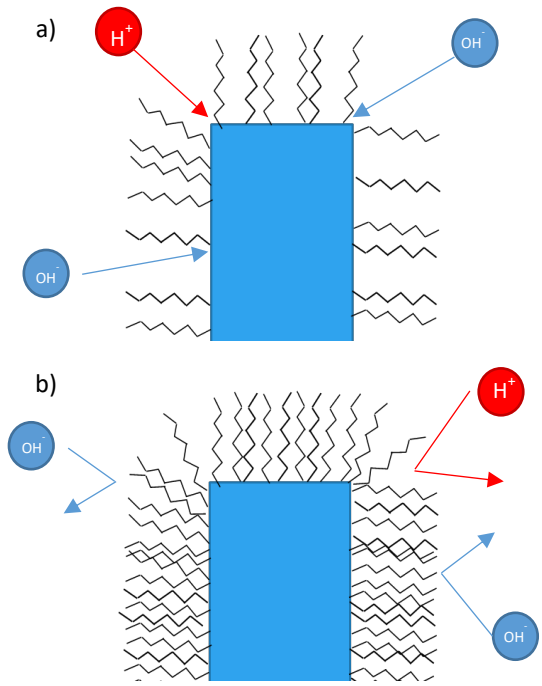


Figure 8 Illustration of the attack of H+ and OH- ions on a coated surface with a) poor b) good surface coverage

Mechanical Durability Testing

Following the mechanical abrasion tests using magnetic stirrers, there is a substantial decrease in the measured contact angles for the paints on all substrates (Figure 9). The nano- and microstructures of the superhydrophobic surface were destroyed by abrasion, leading to a loss in superhydrophobicity. Glass is shown to be the most durable, followed by sandpaper, then aluminium and copper.

A significant decrease in contact angles measured on the aluminium and copper surfaces was observed, which is clearly reflected by the loss of coating during abrasion (Figure 10). The poor adhesion of the coating to the metal surfaces could be attributed to the fact that the coating was applied to a metal oxide layer, which has 5-10 times lower surface free energy than its metal counterparts (Castle, 2008). This leads to a less negative free energy of adhesion between the metal substrate and coating, thereby decreasing interaction stability (Madeira et al., 2018). The hydrophobicity of the metal substrates was tested prior to coating (Figure 9) and shown to be hydrophilic. It was observed that the contact angles after abrasions with metal substrates are slightly higher than coating remained. This could mean that there were paint particles that remained on the metal surfaces, or that the abrasion has resulted in an increase in roughness of the surface itself.

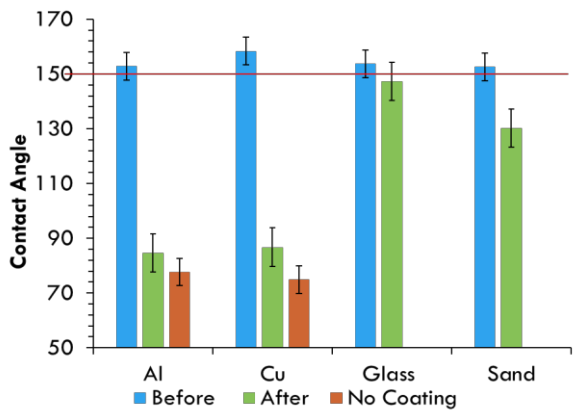


Figure 9 WCAs of coated substrates before and after mechanical durability testing, and for uncoated metals

Conversely, sandpaper and glass were able to retain the coating relatively well during abrasion, even though superhydrophobicity was not maintained for sandpaper. The abrasive minerals on the sandpaper used were aluminium oxide crystals, therefore a low adhesion of the coating to the surface was expected, similar to that of the aluminium substrate. However, due to the rougher topology of the surface, the coating is preserved in between the minerals on the sandpaper, thus maintaining some hydrophobicity. A decrease in contact angle is seen due to the exposure of some of the aluminium oxide, which was found to exhibit hydrophilicity.

For glass, there is no significant change in the coating or the contact angle before and after mechanical testing, especially when taking into its standard deviation. The higher surface energy of glass compared to the metal oxides (Mattox, 2010) results in stronger adhesion of the coating to the surface, preventing the entire coating being abraded off the surface. Improved adhesion between the coating and glass substrate can also be attributed to the homogenously flat surface of the glass, compared to the grooved metals. The flatter surface results in greater wetting of the substrate by the polymeric film, thus resulting in greater interactions

between coating and substrate. Moreover, the flat surface allows for a denser and more efficient packing of particles on the glass leading to greater mechanical strength. The implication of this is that application of the film to flatter substrates may result in improved mechanical durability.

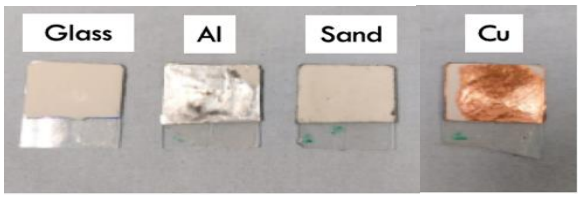


Figure 10 Coated substrates (glass, aluminium, sandpaper and copper) after mechanical durability testing

The maintained superhydrophobicity of the coating on glass could imply volumetric superhydrophobicity, where a good distribution of hydrophobic particles in the paint results in hydrophobicity throughout the coating, not just the surface layer. As the top layers of the coating are abraded, the newly exposed surface also exhibits superhydrophobicity. The even particle distribution is representative of a low Peclet number. Peclet number is defined by the ratio of the rate of drying to the rate of diffusion, with high Peclet numbers resulting in particles being drawn up through the paint by the evaporating solvent and accumulating at the surface (Narita et al., 2004). Volumetric superhydrophobicity is desirable in coatings as the property is maintained even after abrasions. Future work would involve the measurements of drying rate and diffusivities of hydrophobic particles in the paint, in order to determine Peclet numbers and verify the extent to which the paint demonstrates volumetric superhydrophobicity on different substrates. This could help to provide insight into how the film formation of the paint on different substrates differ, and thus how the topology of the substrate surfaces affects this.

Effect of pre-treatment methods on superhydrophobicity

Eggshells are mainly composed of calcium carbonate (96%), with the remainder being a protein matrix (2%) as well as other trace metals such as magnesium and phosphorus (Hincke et al., 2012). Therefore, it was investigated whether the removal of these protein and other impurities by washing with sodium hydroxide would lead to a higher superhydrophobicity, as it was speculated that the presence of these proteins could interfere with the functionalisation with fatty acid (Causby, 2019).

From the results in Figure 11, pre-treatment of eggshells was generally detrimental to the hydrophobicity, displaying a lower contact angle compared to the untreated eggshell coating for all of the pre-treated glass slides. This can be attributed to several reasons. Firstly, it has been shown that sodium hydroxide can attack calcite crystals reducing its size (Simons, Tyler & Thomas, 1966). As the particle size plays a big role in creating the micro- and nano-structures on the surface, necessary for superhydrophobicity, smaller sized particles may produce grooves that are too small to trap air pockets, leading to a lower contact angle. Secondly, following deproteinisation of eggshell powder, not all the sodium hydroxide may have been washed away. During the functionalisation of the eggshell powder with stearic acid, stearic acid may react with the sodium hydroxide residues instead. Therefore, this may result in a reduction in adsorbed stearic acid, leaving some of the calcium carbonate unfunctionalised, thus resulting in lower hydrophobicity. Lastly, calcium carbonate is more soluble in aqueous sodium hydroxide solution as compared to water (Hocking, 2005). This means that pre-treating in sodium hydroxide may lead to the dissolution of some of the calcium carbonate or lead to precipitation of calcium hydroxide on the surface. SEM imaging can be carried out to verify the speculated mechanisms.

Different cleaning procedures of glass slide also were explored to remove oil residues and organic matter from the glass substrates, ensuring cleanliness prior to coating. This ensures that there is good adhesion of the coating to the surface and no pinholes are formed in the coating. The results in Figure 11, has shown that the untreated slide was just as effective in terms of the contact angle as a cleaned slide. As the coating was thick, the cleaning of the glass substrate would not play a big role in the hydrophobicity. However, it may have an impact during mechanical durability tests as the adhesion between the coating and substrates is important. Therefore, for subsequent experiments, the eggshells will not be deproteinised with sodium hydroxide, and no pre-treatment of substrates will be utilised. However, there is evident potential to conduct further investigations into the effect that this has on the mechanical durability of the substrates. Although the literature has shown that acetone and ethanol should have no effect on the surface energies of the substrates, other reagents can be explored to strengthen the interactions between the substrate and paint.

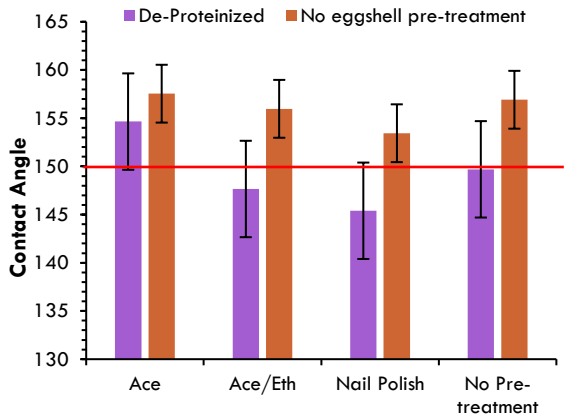


Figure 11 WCAs of eggshell coated glass slides with/without deproteinisation and different slide pre-treatment

Conclusions

This study establishes a facile, environmentally friendly and inexpensive method to fabricate superhydrophobic surfaces with hierarchical structures through pipette coating of paint made from eggshell waste, stearic acid, polyethylene and toluene, based off previous studies by Causby (2019).

The coating exhibited excellent universality on non-porous substrates. However, it was not suitable for coating porous substrates such as filter paper as the coating is drawn into the pores producing an uneven coating on the surface. Pipette coating was the most effective coating method, producing a thick and even coating which displayed the highest contact angles for all substrates. Results from the chemical stability tests revealed that the eggshell coating shows some resistance to various acidic and basic environments in the pH range of 3-11. It should be noted that superhydrophobicity is lost after only 15-minute immersion in all pH solutions which can be due to the dissolution of stearic acid, or the reaction of calcium carbonate and calcium stearate with acidic/basic solutions. However, there was not a significant change in the contact angle after immersions which indicates slow reaction kinetics or good surface coverage of the coating. Mechanical durability tests showed that the adhesive force between the coating and substrate surface may affect long-term stability, which depends on the substrate surface morphology and surface energy. It was found that pre-treatment of glass substrates with acetone and ethanol had no effect on the superhydrophobicity of the paint, implying that there was little effect on the coated surface topology. Deproteinisation of the eggshells with sodium hydroxide negatively affected the hydrophobicity of the paints produced.

This work proposes a further investigation into film formation and drying mechanisms of the superhydrophobic paint and how it influences the surface topology and interfacial interactions between paint and the substrate. Scanning electron microscopy (SEM) and thickness measurements can be used to better characterise and compare the surface morphology and particle distribution of different coated substrates, coating application methods and the surfaces before and after durability tests. Chemical characterization methods such as ion chromatography can be used to further understand and describe the reactions that occur during chemical stability tests. Spray coating and other application techniques can also be explored. The fabrication technology can then be scaled up to be used in the industry.

Acknowledgements

We would like to express our gratitude to Dr Vikram A. Karde for his valuable words of wisdom, continuous support and feedback he gave throughout the project. We also thank all members of the Heng group for their kind help, support and guidance during our time in the laboratory.

References

Alsoy, S. & Duda, J.L. (1998) Drying of Solvent Coated Polymer Film. *Drying Technology*, 16:1-2, 15-44, Available from: doi: 10.1080/07373939808917390

American Elements. (n.d.) Sodium Stearate. Available from: https://www.americanelements.com/sodium-stearate-822-16-2 [Accessed 11th December 2019]

Castle, J.E. (2008) The Composition of Metal Surfaces After Atmospheric Exposure: An Historical Perspective. *The Journal of Adhesion*, 84:4, 368-388. Available from: doi: 10.1080no/00218460802004477

Causby, A.T. (2019) Investigating the formulation and efficiency of antimicrobial coatings derived from waste shells. Imperial College London.

Cree, D. & Rutter, A. (2015) Sustainable Bio-Inspired Limestone Eggshell Powder for Potential Industrialized Applications. *ACS Sustainable Chem*. Eng. 3, 5, pp. 941-949. Available from: doi: 10.1021/acssuschemeng.5b00035

He J., He. J., Yuan M., Xue M., Ma X., Hou L., Zhang T., Liu X. & Qu M. (2018) Facile Fabrication of Eco-Friendly Durable Superhydrophobic Material from Eggshell with Oil/Water Separation Property. Available from: doi:10.1002/adem.201701180

Hincke, M.T., Nys, Y., Gautron, J., Mann, K., Rodriguez-Navarro, A.B. & McKee, M.D. (2012) The eggshell: structure, composition and mineralization. Frontiers in Bioscience 17, 1266-1280. Available from: https://www.researchgate.net/publication/51895246_The_eggshell_structure_composition_and_mineralization [Accessed 1st December 2019]

Hocking, M.B. (2005) '7 – Industrial Bases by Chemical Routes' in Handbook of Chemical Technology and Pollution Control (3rd Edition). Academic Press, pp. 201-220.

Jafari, R. & Farzaneh, M. (2012) Development a simple method to create the superhydrophobic composite coatings. *Journal of Composite Materials*, Volume 47, Issue 25, pp. 3125-3129. Available from: doi: 10.1177/0021998312462429

Jeevahan, J., Chandrasekaran, M. Britto Joseph, G., Durairaj, R.B. & Mageshwaran, G. (2018) Superhydrophobic surfaces: a review on fundamentals, applications, and challenges. *J Coat Technol Res* 15, 231–250. Available from: doi:10.1007/s11998-017-0011-x

Kokkanis, P. (2017) Durability of superhydrophobic coatings from recycled materials. Imperial College London

Madeira, D.M.F., Vieira, O., Pinheiro, L.A. & Carvalho, B.M. (2018) Correlation between Surface Energy and Adhesion Force of Polyethylene/Paperboard: A Predictive Tool for Quality Control in Laminated Packaging. *International Journal of Chemical Engineerin*g, vol. 2018. Available from: doi: 10.1155/2018/2709037

Martin, P.M. (2010) 'Chapter 3 – Surface Preparation for Film and Coating Deposition Processes' in Handbook of Deposition Technologies for Films and Coatings (3rd Edition). William Andrew, pp. 93-134.

Mattox, D.M. (2010) Handbook of Physical Vapor Deposition (PVD) Processing (2nd Edition). William Andrew

Narita, T., Hebraud, P. and Lequeux, F., 2005. Effects of the rate of evaporation and film thickness on nonuniform drying of film-forming concentrated colloidal suspensions. *The European Physical Journal E*, 17(1), pp.69-76.

National Center for Biotechnology Information. (2004) Calcium carbonate. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/Calcium-carbonate [Accessed 7th December 2019]

National Center for Biotechnology Information. (2004) Stearic acid. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/Stearic-acid [Accessed 7th December 2019]

Ossila. (n.d.) Dip Coating Theory: Film Thickness. Available from: https://www.ossila.com/pages/dip-coating-theory-film-thickness [Accessed 11th December 2019]

Queiroz Antonino, R.S.C.M., Lia Fook, B.R.P., Oliveira Lima, V.A., Farias Rached, R.I., Nascimento Lima, E.P., Silva Lima, R.J., Peniche Covas, C.A. & Lia Fook, M.V. (2017) Preparation and Characterization of Chitosan Obtained From Shells of Shrimp *(Litopenaeus vannamei Boone*). *Mar Drugs*, 15(5), 141. Available from: doi: 10.3390/md15050141

Sarkar, A., Ghosh, A.K. & Mahapatra, S. (2012) Lauric acid triggered in situ surface medication and phase selectivity of calcium carbonate: its application as an oil sorbent. *J. Mater. Chem*., 22, 11113. Available from: doi: 10.1039/c2jm30778b

Shi, X., Rosa, R. & Lazzeri, A. (2010) On the Coating of Precipitated Calcium Carbonate with Stearic Acid in Aqueous Medium. *Langmuir*, 26(11), 8464-8482. Available from: doi: 10.1021/la904914h

Simons, P.C.M., Tyler, C. & Thomas, H.P. (2007) The effect of sodium hydroxide and sodium sulphide on the snapping strength of eggshells. *British Poultry Science* 10:2, pp. 165-174. Available from: doi: 10.1080/00071666608415638

Simpson, J.T., Hunter, S.R. & Aytug, T. (2015) Superhydrophobic materials and coatings: a review. *Rep. Prog. Phys*. 78 086501. Available from: doi: 10.1088/0034-4885/78/8/086501

Smith, R., Inomata, H. & Peters, C. (2013) 'Chapter 4 – Historical Backgrounds and Applications in Supercritical Fluid Science and Technology Volume 4.

Tyona, M.D. (2013) A theoritical study on spin coating technique. *Advances in Material Research*, Vol. 2, No. 4, pp. 195-208.

Verma, N., Kumar, V. & Bansal, M.C. (2011) Utilization of Egg Shell Waste in Cellulase Production by Neurospora crassa under Wheat Bran-Based Solid State Fermentation. *Pol. J. Environ. Stud.* Vol. 21, No. 2, pp. 491-497. Available from: http://www.pjoes.com/Utilization-of-Egg-Shell-Waste-in-Cellulase-r-nProduction-by-Neurospora-crassa-under,88777,0,2.html [Accessed 10th December 2019]

Vuckovac, M., Latikka, M., Liu, K., Huhtamaki, T. & Ras, R.H.A. (2019) Uncertainties in contact angle goniometry. *Soft Matter*, 15, 7089. Available from: doi: : 10.1039/c9sm01221d

Wang, Z., Yi, M., Zhang, Z., Guo, M., Lu, P., Chen, Z. & Wang, S. (2017) Fabrication of highly water-repelling paper by surface coating with stearic acid modified calcium carbonate particles and reactive biopolymers. *Journal of Bioresources and Bioproducts*, 2(2), pp. 89-91. Available from: 10.21967/jbb.v2i2.91

Wen, G., Huang, J.X. & Guo, Z.G. (2019) Energy-effective Superhydrophobic Nanocoating Based on Recycled Eggshell. Colloids and Surfaces *A: Physicochemical and Engineering Aspects*, Volume 568, Pages 20-28. Available from: doi: 10.1016/j.colsurfa.2019.01.067

Younes, I. & Rinaudo, M. (2015) Chitin and Chitosan Preparation from Marine Sources. Structure, Properties and Applications. *Mar Drugs*, 13(3), 1133-1174. Available from: doi: 10.3390/md13031133

Zhang H., Zeng X., Gao Y. F., Shi F., Zhang P. & Chen J. G. (2011). A Facile Method to Prepare Superhydrophobic Coatings by Calcium Carbonate. *Ind. Eng. Chem. Res*. 50, 6, 3089-3094. Available from: doi:10.1021/ie102149y

Corrosion-Resistant Sol-Gel Coatings for 304 Stainless-Steel Cookstoves

Adrian Lua Eu Kane & Han Chang Eu
Department of Chemical Engineering, Imperial College London, United Kingdom

Abstract

An experimental study on the characteristics of 304 stainless steel after the application of corrosion resistant sol-gel coatings on its surface. Experiments have been conducted with 150mm X 150 mm X 1.16mm 304 stainless-steel sheets which were polished with silicon carbide sandpaper with increasing grit size and subsequently degreased with acetone, before undergoing an alkali batch with 2.5M sodium hydroxide (NaOH). The titanium-based sol gel coatings were prepared via the Pechini method, with titanium (IV) isopropoxide (TIP) acting as the precursor. To achieve uniform thin layer coating, four-sided film applicator assisted by a coating coater machine was used when applying the gel on the surface of the 304 stainless-steel sheets, which emulated the doctor blade coating method. Once applied, the gel was dried in an oven at 100°C for an hour before calcining it at 500°C with a 6°C/min ramp and a dwelling time of 5 minutes. Raman spectroscopy analysis was conducted to confirm the presence of titanium dioxide (TiO_2) on the surface of the sheet. Contact angle analysis was also performed to determine the wettability of the sol-gel on the stainless-steel sheet. Multiple layers of coating were applied to increase the density of TiO_2 on the surface of the stainless-steel sheet, hence improving its corrosion-resistant capabilities. Combustion tests proved successful, with the coatings remaining after intense burning under direct fire fuelled by firewood.

Keywords: 304 stainless-steel, sol-gel coating, anatase TiO_2, Pechini method, Raman spectroscopy

1. Introduction

Approximately of 777 million people in India still rely on solid biomass fuel, with 77% of households using firewood specifically, which releases a large amount of carbon monoxide and soot due to incomplete combustion, causing severe health and environmental damages to the region. (Energy, 2019) (Bloomfield, 2014) Many social problems in India can also be attributed to the inefficient burning of solid biomass cooking fuels, especially among women and children, who spend most of their time gathering biomass fuel rather than pursing activities such as education or income generation. (Brady, et al., 2017)

With the sole intention of solving this complex problem, Prakti, a social enterprise dedicated to disrupting decades of biomass cookstove design has taken significant strides in to ensuring that their products are efficient, cheap and easy to use. (Prakti, 2019)

Although Prakti cookstoves are significantly more efficient than traditional open fire fuel combustion, there are still serious threats of the cookstoves being vulnerable to pitting corrosion induced by high temperatures and inorganic salts from the firewood which could potentially compromise its integrity and shorten its lifespan. Complementing the research that has already been done by team in Prakti, this paper further investigates the development of a simple and cheap process for applying ceramic coatings to Prakti cookstoves for both corrosion protection and to increase the rate of oxidation of harmful by-products, lowering emissions and increasing the quality of Prakti products.

Figure 1: Cooking condition in India's households *(Energy, 2019)*

Figure 2: Example of a standard Prakti cookstove *(Prakti, 2019)*

Background

1.1 Sol-gel and Pechini Method

Sol-gel is widely defined as a wet chemical process that involves the formation of an inorganic colloidal suspension (sol) and the

gelation of the sol in a continuous liquid phase (gel), forming a three-dimensional network structure. These 'sols' are formed when starting materials, usually metal organic compounds such as metal alkoxide or inorganic metal salts goes through a series of hydrolysis and polymerization process, resulting in the formation of the said colloidal suspension. (Sajjadi, 2005) (Ullattil & Periyat, 2017)

Due to its wide range of applications, including protective coatings, catalysts and insulating materials (Gelest, Inc, 2019), extensive research has been done on the subject dating as far back as the year 1845, when Ebelman first observed the formation of a transparent glass following exposure to the atmosphere of a silane obtained from silicon tetrachloride ($SiCl_4$) and ethanol (Dimitriev, et al., 2008). This was further extended by Geffcken and Berger from Shott Company in 1930, who pioneered the way of achieving a sol-gel process from oxide layers on industrial glasses using combinations of metal-containing precursors by deep coating (Gelest, Inc, 2019) (Dimitriev, et al., 2008). Following that, work from H. Schroeder, who deposited transparent coatings on glass surfaces in order to correct the refractive index using titanium butoxide and Dislich, who elaborated the chemical basis for the preparation of multicomponent oxides and ceramics also contributed significantly to the rapid development of sol-gel technology (Dimitriev, et al., 2008) (Dislich & Hinz, 1982).

Over the years, the Pechini method, an alternative to the sol-gel method, has gain more popularity due to the simplicity and versatility of the process (Sunde, et al., n.d.) (Dimesso, 2018). The general principles of the process involves the formation of a chelate between mixed cations with a hydroxycarboxylic acid with the aid of poly-alcohols. Gels are form when the chelates cross-linked through esterification (Sunde, et al., n.d.) (Dimesso, 2018) (Xavier, 2004). calcination at high temperature, a well-structured metal oxide lattice is formed, which can be used to provide a protective and corrosion resistant layer on metal surfaces.

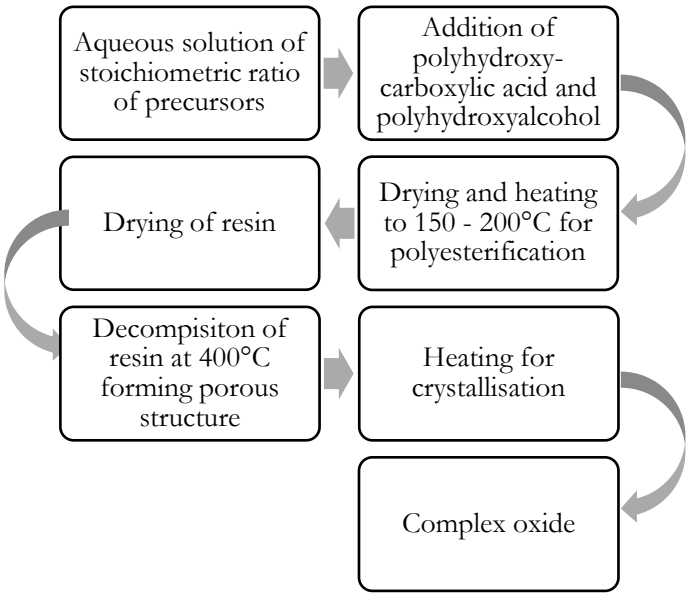


Figure 4: Flowsheet of the modified Pechini process *(Dimesso, 2018)*

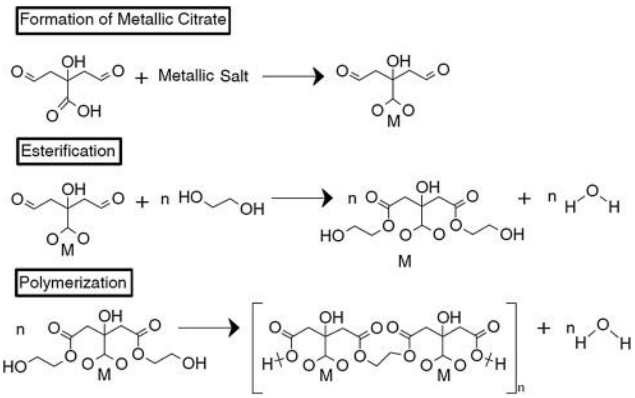


Figure 3: Schematic representation of the Pechini method *(Xavier, 2004)*

2.2 Gel Coating

Gel coating on a surface are predominantly achieved via dip or spin coating, with the former being widely considered as the easiest method to execute (Sunde, et al., n.d.). The general principle of this process involves the withdrawal of the substrate, initially dipped in the gel, at a constant speed (Kumar, et al., 2016). The gel homogenously stretches across the surface of the substrate due to the mutal effect of capillary rise and viscous pull, both of which are sensitive to the withdrawal speed (Kumar, et al., 2016) (Ainuddin & Aziz, 2016).

In comparison to dip coating, the spin coating method also involves the thining of the gel coating via evaporation and centrifugal draining (Kumar, et al., 2016) (Brinker, et al., 1992). Excess amount of gel are expelled from the surface of the substrate, retaining the uniformity of the coating until any variation in the viscosity occurred. Again, this process is sensitive to the rate of rotation (Çağlar, et al., 2008).

Keeping in mind the need for consistent and repeatable results, this paper looks to investigate further into the use of the doctor blade method to achieve even gel coating on the surface of the metal. In this process, the gel is poured on to the surface and allowed to spread evenly acorss the surface when a constant relative movement is established between the blade and the surface (Berni, et al., 2004). As the coatings achieved are only micrometre-thin, the effect of multi-layer coating has also been explored as a way to increase the durability and intergity of the coating.

2.3 Gel Drying and Calcination

It is frequently reported that the process of drying plays an integral role in achieving successful gel coating. After coating of the metal surface is completed, the drying process evaporates the solvent and remove organic residuals so as to not compromise the corrosion-resistant capabilities of the gel (Çağlar, et al., 2008). Common methods of drying includes drying under room conditions or high temperature in an oven (Xavier, 2004) (Ainuddin & Aziz, 2016) (Brinker, et al., 1992). After drying, different calcinining temperatures and rate of temperature increase can be utilised to affect the properties of the ceramic coating, which in turn can contibute towards the life span of the coated metal (Ainuddin & Aziz, 2016) (Liu, et al., 2003).

2.4 *Metal Surface Pre-Treatment*

Different methods of pretreating the metal surface have been reported which includes the use of silicon carbide abrasive discs, alumina solution and diamond paste (Ainuddin & Aziz, 2016) (Ćurković, et al., n.d.) (Shen, et al., 2005). Alternative methods such as alkaline etching have also been previously looked at and will continue to be further explored in this paper (Balamurugan, et al., n.d.). Post-treatement, the metal surface is ultrasonically cleaned with acetone.

2.5 *Corrosion*

Corrosion can be described as the degradation of materials by spontaeous reaction with their environments. It is predominantly an electrochemical process which can be influenced by chemical, physical and mechanical factors. Firewood contains inorganic elements such as chlorine which exist either as hydrochloric acid (HCl) or condensed chlorides (e.g. NaCl). As shown in Table 1, the mechanism of the pitting corrosion begins with the formation of chlorine gas from the oxidation of HCl or NaCl reacting with the oxide scales (i.e. Fe_2O_3 and Cr_2O_3) and oxygen. When chlorine penetrates the oxide scale via cracks or pores, solid metal chlorides (i.e. $FeCl_2$) at the metal/scale interface. At temperatures around 500°C, $FeCl_2$ will volatise into a gaseous state then oxidise to form the metal oxide (i.e. Fe_2O_3 and Fe_3O_4) and chlorine. This mechanism is known as active oxidation by chlorine and it is self-sustaining (Antunes & De Oliveira, 2013).

Table 1: Corrosion mechanism due to wood burning *(Antunes & De Oliveira, 2013)*

$2HCl + 1/2O_2 \leftrightarrow Cl_2 + H_2O$	**(1)**
$2NaCl + Fe_2O_3 + 1/2O_2 \leftrightarrow Na_2Fe_2O_4 + Cl_2$	**(2)**
$4NaCl + Cr_2O_3 + 5/2\ O_2 \leftrightarrow 2NaCrO_4 + 2Cl_2$	**(3)**
$Fe + Cl_2 \leftrightarrow FeCl_2(s)$	**(4)**
$FeCl_2(s) \leftrightarrow FeCl_2(g)$	**(5)**
$3FeCl_2 + 2O_2 \leftrightarrow Fe_3O_4 + 3Cl_2$	**(6)**
$2FeCl_2 + 3/2O_2 \leftrightarrow Fe_2O_3 + 2Cl_2$	**(7)**

3. Methods

3.1 *Metal Preparation*

All 304 stainless-steel sheet substrates used in the experiment have dimensions of 150 mm X 150 mm X 1.16 mm. These substrates were then quartered into 75 mm X 75 mm X 1.16 mm coupons before undergoing the bending test and combustion tests with firewood.

3.2 *Gel Preparation*

The 304 stainless-steel sheets were coated with a titanium-based gel by utilising the Pechini method, with titanium (IV) isopropoxide (TIP) acting as the precursor. Before TIP can be added, ethylene glycol (EG) was first heated up to 65°C on a uniform heat source. Whilst the TIP and EG were undergoing vigorous mixing with the use of a magnetic stirrer, citric acid (CA) powder was slowly added into the solution at a total molar ratio of EG:CA:TIP of 1:0.26:0.028. It was essential for this process to take place slowly as accelerating reactions through short times causes precipitates to form instead of a uniform gel network, subsequently resulting in the solution turning cloudy and negatively affecting the gel's properties. Following this, the temperature was raised to 90°C for 2 hours to allow for complete homogenous dispersion of titanium ions across the gel polymer network. The temperature was then subsequently lowered to 60°C for half an hour before the sol-gel was then cooled at room temperature. The resulting gel is yellow in colour, largely contributed by the TiO2 present in the gel.

3.3 *Surface Pre-Treatment*

Two distinct pre-treatment methods were investigated to understand the effect of pre-treatment on the overall interaction between the stainless-steel sheet and the gel. The sheets were first polished with silicon carbide (SiC) sandpaper starting from 600 grit size, followed by 800 and finally 1200 grit size, removing the pre-existing oxide layers on the surface of the sheet. For the 'only-polished' (OP) method, the substrates were only further rinsed with acetone to remove any grease on the surface. For the 'polished with alkali bath' (PBA) method, the substrates were submerged in 2.5M NaOH solution at 90°C for 15 minutes post-degreasing to allow for chemical etching. Following that, the substrates were rinsed with deionised water and dried using instrument air to remove any residual contaminants on the surface. Untreated control substrates were also prepared by just degreasing the surface with acetone with no other pre-treatment steps performed.

3.4 *Coating*

The doctor blade coating method that was simulated for this experiment involves the use of a four-sided film applicator and control coater machine. Firstly, the 304 stainless-steel sheet substrates were placed on top of a smooth, flat surface. The 10 µm side of the SZQ four-sided wet film coating applicator No.1 was placed on the substrate surface before a small amount of sol-gel was poured on to substrate surface in front of the applicator. The temperature of the sol-gel should be maintained at 60°C before performing the coating procedure. A bar, which was attached to a K Control Coater machine, set at an average travelling speed of 0.04 ms^{-1}, was utilised to slowly push the film applicator forward in a single stroke, resulting in a thin layer coating on the substrate surface. Excess gel was wiped and cleaned with deionised water. Figure 5 demonstrates how the set-up was performed.

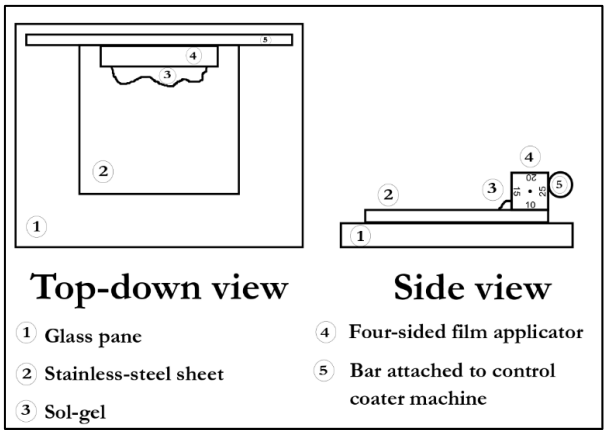


Figure 5: Doctor blade method set-up

3.5 Drying

The 304 stainless-steel sheet substrates were dried under a pressure of -800mbar using Technico OVA031.XX3.5 vacuum oven. The substrates were dried at temperatures at 100°C for an hour to remove water molecules present in the sol-gel and form a resin.

3.6 Calcination

After the drying process, the substrates were then immediately transferred to a Carbolite CWF 1200 furnace for calcination. Due to the enhanced corrosion resistance capabilities of anatase form of TiO_2, the substrates were calcined at a temperature of 500°C with a ramp speed of 6°C/min and a dwelling time of 5 minutes. It was also important to ensure that the ramp speed falls between 5-10°C/min as an excessive increase in the rate of temperature rise will cause cracking and flaking of the surface. The steel sheets were then removed from the furnace and allowed to cool at room temperature post-calcination.

3.7 Multi-Layer Coating

To assess the effect and effectiveness of multi-layered coatings, repeated coatings were conducted using the doctor blade method. Subsequent coatings can only be applied after the previous coating has been calcinated and allowed to cool at room temperature. The substrate surface was first rinsed thoroughly with deionised water before all the steps from Section 3.4 onwards were repeated. Analysis was conducted on substrates that has been coated with the titanium-based sol gel three and five times.

3.8 Surface Analysis

Raman spectroscopy was performed with the SENTERRA II Compact Raman Microscope to determine the presence of the TiO_2 on the surface of the substrates. A laser wavelength of 532 nm and power of 12.5 mW was used alongside a resolution of 1.5 cm^{-1}, spectral range from 50 cm^{-1}–1450 cm^{-1}, coadditions set to 1 and integration time of 6 seconds. The SENTERRA II also features an optical microscope with 10X magnification which was used to take images of the coated surfaces of the substrates. Contact angle analysis was performed with the Krüss DSA30 Drop Shape Analyser. This step involved placing a drop , equivalent to about 0.3 μL, of the sol-gel which was pre-heated to 60°C beforehand onto the surface of the substrates which has been cleared of any impurities and the contract angle between the drop and the surface is then measured and recorded by the instrument for a period of 30 minutes.

3.9 Bending Test

The samples were quartered from their original size into coupons with the size of 75 mm X 75 mm X 1.16 mm. The coupons are then bent to an angle of 90° with the coated layer facing inwards. The surface along the bent region is then observed under an optical microscope at 10X magnification to observe for any cracks or deformities.

3.10 Combustion Test

The substrate coupons also underwent a combustion test where the samples are placed in a firewood stove fuelled by UK hardwood two season dried to simulate real life usage. The test was split into 3 cycles lasting 6 hours each, totalling to 18 hours. Table 2 describes how the substrates were arranged and placed in the stove for each cycle.

Table 2: Arrangement and position of samples in stove

Cycle	Arrangement and position in stove
1	Samples were placed under the wood at the back of the oven and arranged in single file.
2	Same position in the oven as Cycle 1 but the samples have swapped places with their neighbouring samples.
3	Same arrangement as Cycle 1 but now the samples have been brought to the front of the stove where it was closer to the fire.

4. Results & Discussions

4.1 Surface Pre-treatment

The contact angle analysis was used to quantify the wettability of the gel on the surface of the substrate, which will then be used to compare between the two different pre-treatment methods. In general, angles that are greater than 90° indicates poor wetting of the liquid phase on the surface of the solid, while angles lesser than 90° would indicate the opposite where there is good wettability of the liquid phase with the solid surface.

For the control sample as well as the OP and PBA substrates, the gel exhibited good wettability since all three of them have contract angles which were less than 90°. Among the three, the PBA substrate had the lowest contact angle after 30 minutes compared to the other two, despite having the highest starting contact angle. This can be attributed to several factors. Firstly, the polishing of the substrate with SiC sandpaper increased the surface roughness which contributed to lower contact angles and better wettability (Anon., 2019). The chemical etching step performed on the PBA substrate as, described back in Section 3.3, had formed a high surface density of Fe-OH bonds which had improved the interaction between the metal surface and gel (Ansari, et al., 2015).

The premise of pre-heating the metal substrate to 60°C before performing the contact angle analysis to improve the gel's wettability was also investigated. However, based on the results obtained as shown in Table 5, there was no improvement in the contact angle, which was attributed to the substrate cooling down too fast back to ambient temperature and the lack of viable methods to maintain the substrate temperature at 60°C throughout the duration of the analysis.

Table 3: Contact angle at minute 0 and 30

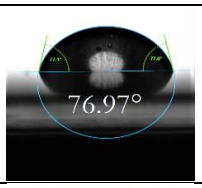
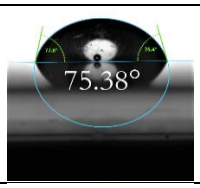
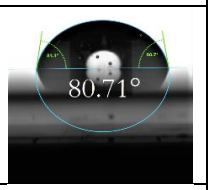
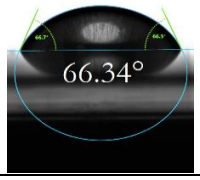
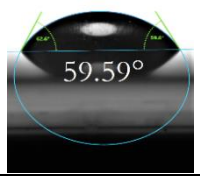
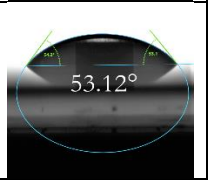
	Control	OP	PBA
Minute 0			
Minute 30			

Table 4: 10X magnification on substrate surface

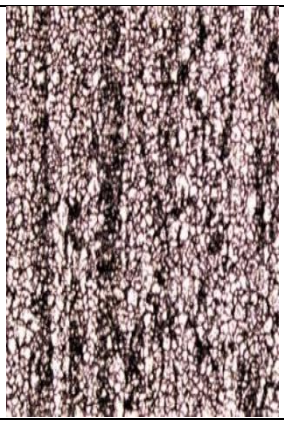
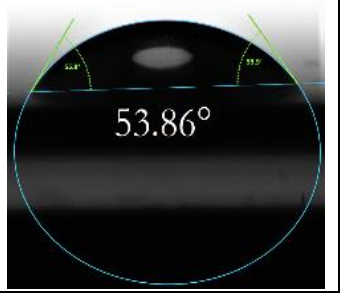
Control substrate	Substrate polished with SiC sandpaper
	

Table 5: Contact angle at minute 30 for unheated PBA and pre-heated PBA

Unheated PBA	Pre-heated PBA
	

4.2 Coating

The doctor blade coating method was employed as a means of obtaining a coating of uniform thickness consistently. This method also has an advantage where the coating can be done at a faster rate than the dip coating method. It is important to note that the sol-gel should be maintained at 60°C to reduce the viscoelasticity of the sol-gel as shown in Figure 6. With a reduced viscoelasticity, this would allow the sol-gel to spread better on the surface of the substrate and there would be less shrinkage of the are covered by the sol-gel. Figure 7 shows the contact angles obtained throughout the 30 minutes.

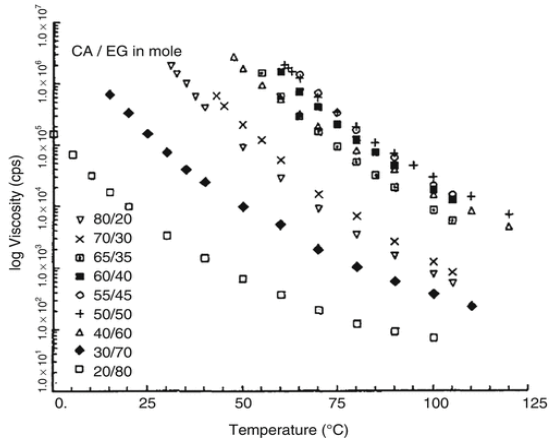


Figure 6: Graph of viscosity of sol-gel with different CA:EG ratios versus temperature (Sunde, et al., n.d.)

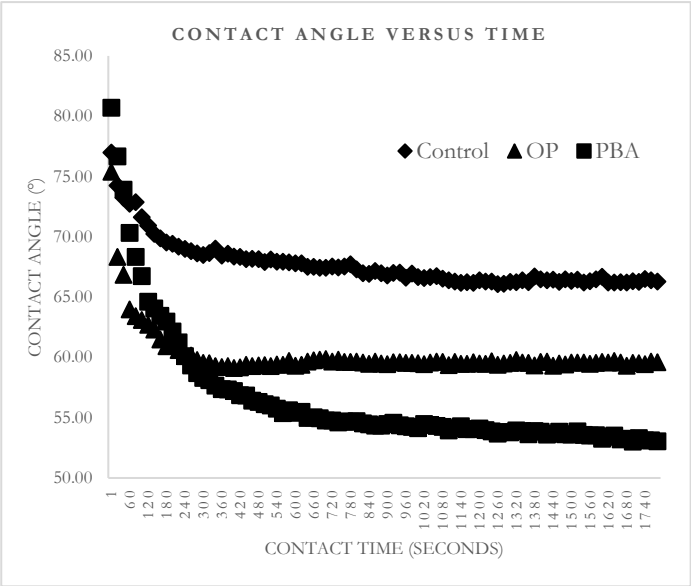


Figure 7: Graph of contact angle versus time

The 10 µm side of the four-sided film applicator was utilised instead of the 15 µm, 20 µm or 25 µm side was due to thicker gel layers having a higher tendency to convert into a powder state during the calcination process and not leave any coating on the surface. It can be hypothesised that increased layer thickness would lead to more water molecules persisting in the sol-gel polymer network even after the drying process due to increased mass and heat transfer resistance. These remaining water molecules would act as impurities which disrupts the arrangement of the TiO_2 molecules, resulting in a less cohesive agglomeration and hence increasing the chances of turning into powder when calcinated.

4.3 Drying

A drying temperature of 100°C was selected as it was the highest temperature which allowed for drying of the gel layer on the substrate surface without altering the physical properties (i.e. colour) of the sol-gel. As for the drying time, one hour was found to be sufficient at 100°C as further drying times yield no observable changes in the physical properties.

After the drying process has been completed, the gel converts into a resin-like form where the viscosity of the gel increases due to the evaporation of the water molecules from the gel. An issue that arises during this process was the surface of the resin becoming non-uniform shown in Figure 8. It was observed that regions closer to the edges of the coated area would have a greater resin thickness compared to the regions nearer towards the centre. This can be attributed to the uneven distribution of water molecules in the polymer network which would lead to water molecules being evaporated at different rates across the sol-gel surface.

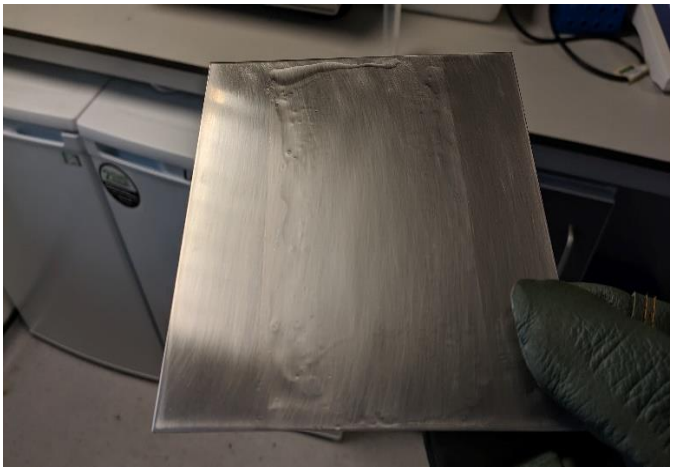


Figure 8: Non-uniformity of the resin surface

4.4 Calcination

After the calcination process, a coating with iridescent effects was formed as shown in Figure 9, which cannot be washed or rubbed off from the substrate surface. The coating can still be scratched with metallic objects (e.g. keys) but would require a significant amount of forced to do so. It was observed that not all the initial coated area features the coating, especially at regions closer to the edges. As discussed in the previous section, the unevenness of the resin surface would mean that some regions of the resin would be thicker than the others, hence increasing the chances that it will turn into powder during the calcination process and not adhere to the substrate surface.

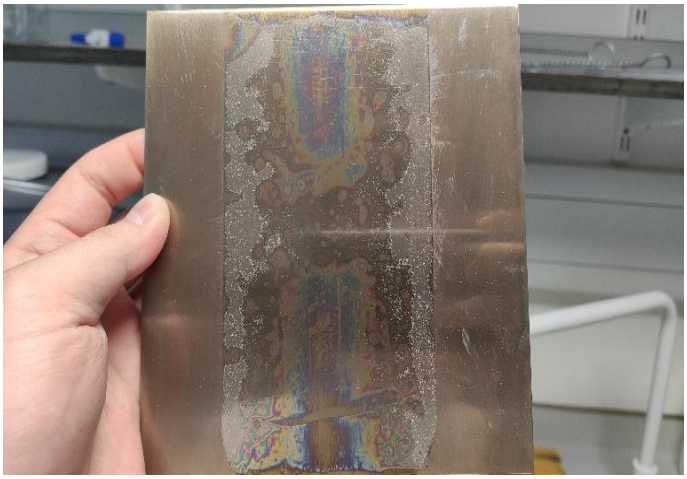


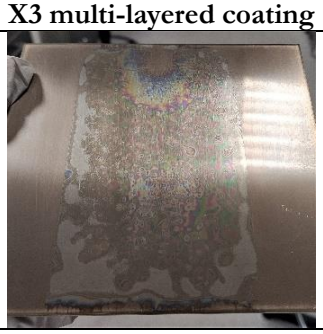
Figure 9: TiO_2 coating formed after calcination process with uncoated regions present at the edges

The calcination temperature of 500°C was utilised in order to obtain TiO_2 in its anatase form, the reason being anatase having a higher corrosion resistance due to the formation of crystallised oxide layer barrier (Kubiak, et al., 2011). A lower calcination temperature would also mean lower energy usage which translates to savings in terms of the cost to perform this coating technique. As for the temperature ramp speed of 6°C/min, a low ramp speed would result in a less cracked coating layer (Liu, et al., 2003), meaning that the coating would be less compromised when acting as a corrosion barrier. A dwelling time of 5 minutes provided a nice balance as a lower dwelling time would not be enough to eliminate all of the organic materials which would leave burnt residue on the coating while a higher dwelling time would increase the likelihood of the sol-gel turning into powder. A higher dwelling time would also increase the cost of this process.

4.5 Multi-Layer Coating

There are two main motivations for applying multiple layers of coating. The first was to reduce the surface area that has been left uncoated and the second being to increase the density of anatase TiO_2 for better corrosion resistance. Even after applying the third layer of coating, there were still uncoated areas which indicates that a higher number of layers would be required. It was only the fifth layer of coating where there is almost complete coverage by the coating. The most obvious downside to this procedure is that it will cost more time and money, but the benefits provided is more than enough to justify it.

Table 6: Comparison between X3 and X5 multi-layered coating

X3 multi-layered coating	X5 multi-layered coating
	

4.6 Raman Spectroscopy Analysis

The Raman spectroscopy analysis not only would confirm the presence of anatase TiO_2 on the surface of the coated substrates but would also help to characterise the effectiveness of applying multiple layers of coating. From the Raman spectrum for anatase TiO_2 obtained from Stagi et al. (Stagi, et al., 2014), the main peaks that can be observed would 2 E_g peaks at 144 cm^{-1} and 639 cm^{-1}, a B_{1g} peak at 399 cm^{-1} and a A_{1g}/B_{1g} doublet peak at 516 cm^{-1}. In total, this would mean that there are 6 active Raman Modes (1 A_{1g} + 2 B_{1g} + 3E_g).

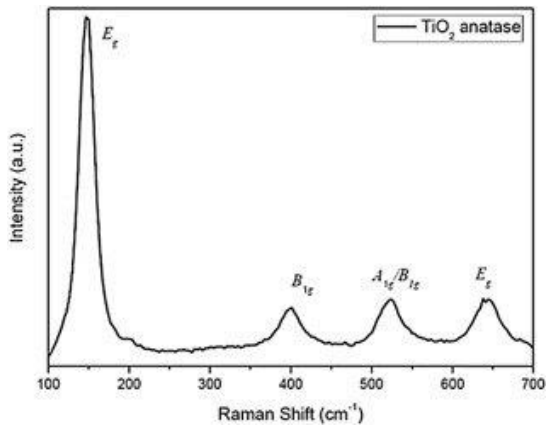


Figure 10: Raman spectrum for anatase TiO_2 (Stagi, et al., 2014)

The same Raman spectrum can be observed for all types of the samples (OP X3 and X5, PBA X3 and X5) when the coated areas were analysed. In terms of the height of the peaks, the PBA X5 featured the highest peaks among the 4 types of samples, followed by OP X5, PBA X3 and OP X3. Hence, the surface of PBA X5 samples has the highest density of anatase TiO_2 and OP X3 samples have the lowest. When comparing the peak heights of PBA X5 and PBA X3 samples, PBA X5 is about 45% taller than PBA X3, indicating that it was worthwhile performing up to 5 layers of coating.

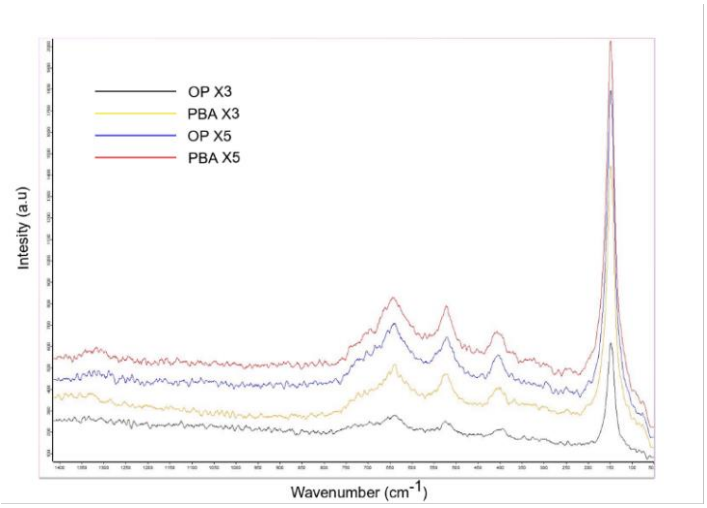


Figure 11: Comparison between OP X3, OP X5, PBA X3 and PBA X5

The surfaces of the coated substrate samples were also placed under the optical microscope at 10X magnification and it was observed that the iridescent effect and the vibrancy of the colours increases from OP X3 and followed by OP X5, PBA X3 and PBA X5. The iridescent effect was caused by multiple reflections of visible light from the transparent TiO_2 films, which leads to phase shifting and interference of the reflections (Masuda & Kato, 2009). It would not be unreasonable to infer that increasing the number of layers as well as the interaction between the sol-gel and the substrate surface would increase said effects.

Cracks and pits were also observed, represented by the dark lines on coating layer. These were caused by thermal stress occurred when the samples were cooled from the calcination process. Given the low thickness of the obtained coatings, this factor would have made them more susceptible towards thermal stress.

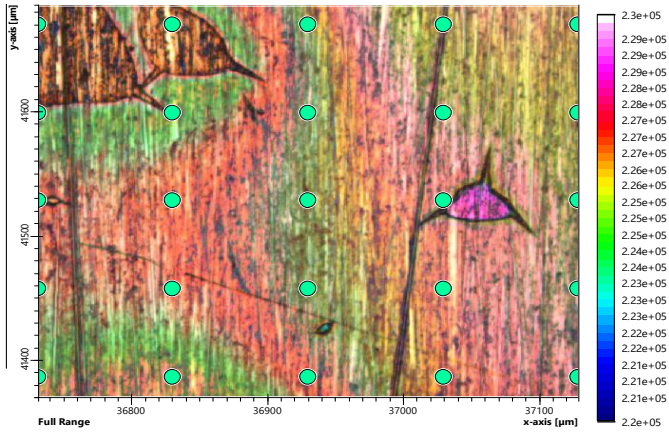


Figure 12: 10X magnification on OP X3 surface

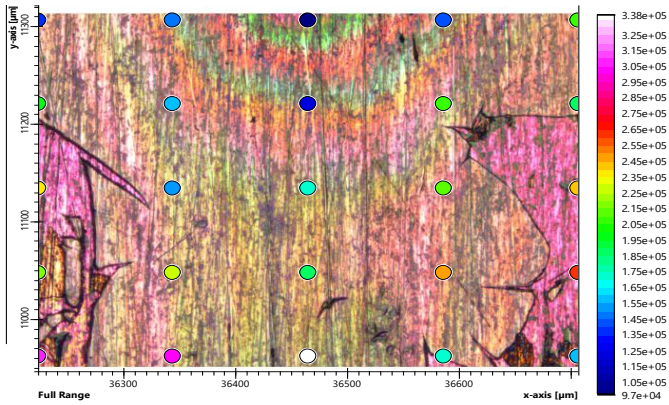


Figure 13: 10X magnification on PBA X3 surface

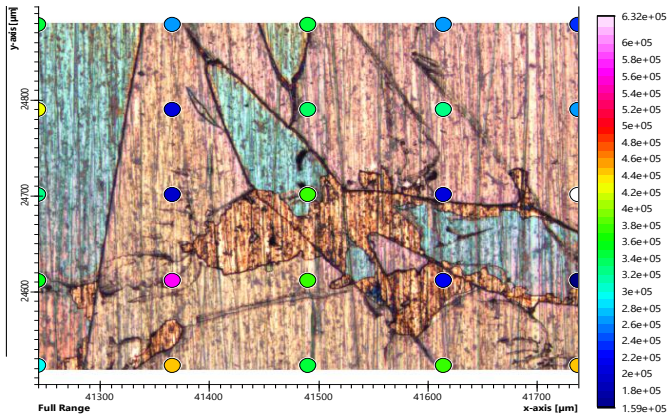


Figure 14: 10X magnification on OP X5 surface

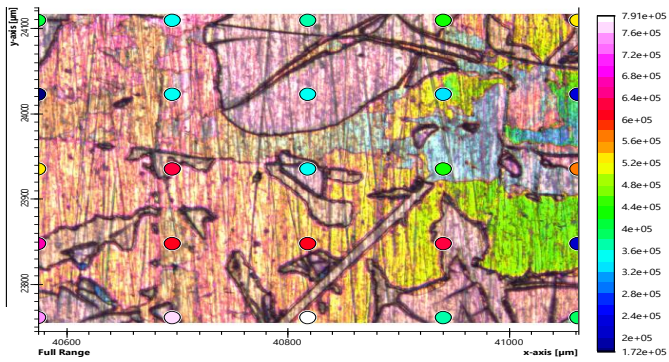


Figure 15: 10X magnification on PBA X5 surface

4.7 *Bending Test*

One of the methods proposed if this coating technique was to be scaled up was coating a large stainless-steel sheet initially, which would then be later cut and bent into the desired shape during the manufacturing process.

Figure 16: Coated substrate after bending test

After the bending test was conducted, it was noted that a small amount of the coating was removed in the form of small specks of flakes along the joint where it was bent, indicating that there are microfractures present. However, no visible cracks can be observed with the naked eye and the coating maintained its adhesion on the surface of the substrate, hence the coating was considered durable against bending stress.

4.8 *Combustion Test*

The 18 hours that spent to perform the combustion test was about equivalent to about 9.5 days' worth of cooking in India (Worldwide & GfK, 2014), hence this can be considered a relatively severe stress test for the samples.

Figure 17: Coated substrate after combustion test (coated region at the top and uncoated at the bottom)

Even after the 18 hours in the stove, the coating seems to have performed quite satisfactorily as it was still adhering to the surface and the coated region experienced less decolourisation and tarnish from the burn residue compared to the untreated part. The burnt samples were also looked through the optical microscope and it was noted that the coated region was cleaner looking than the uncoated region, consistent with what we had observed with the naked eye earlier.

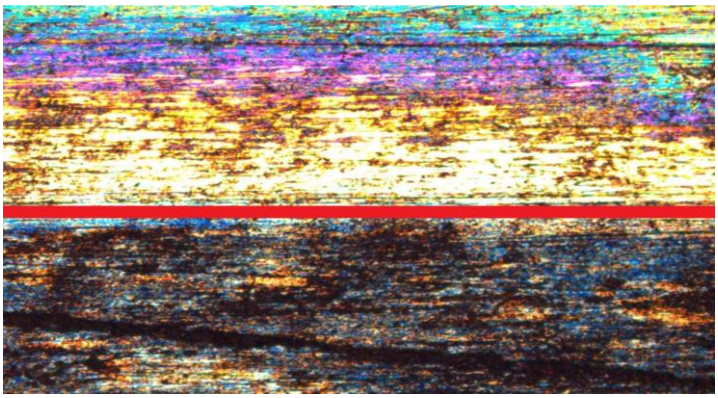


Figure 18: 10X magnification of the border region between the coated area (top) and the uncoated area (bottom)

Despite these promising results, future combustion tests should be performed for a longer period and in controlled conditions that can mimic the ambient environment of India (e.g. high humidity, using firewood native to India) to come to a more conclusive result.

4.9 *Financial Analysis*

Keeping in mind the need to successfully translate the gel synthesis, application and metal pre-treatment process from a lab-scale setting to the large Indian market, a price analysis, based on data obtained from Sigma Aldrich (Sigma Aldrich, n.d.) was conducted to verify the economic feasibility of these processes. Assuming a coating thickness of 10μm, it was calculated that the cost of the coating is at £0.82 (INR76.54) per square metre while the cost of the alkali bath treatment process is also relatively low at £3.21 (INR300) per 1000 millilitre of sodium hydroxide solution used. Since processes will be scaled up further to an industrial level, the cost is expected to be even lower due to the recyclable nature of certain reagents used (mainly sodium hydroxide solution) as well as economies of scale. Taking this into consideration alongside the surface area of the combustion chamber (0.11m^2) and the current cost of a Prakti stove in India (£12 or INR1121.54), it is computed that the total cost for the cookstove will only increase further by approximately 0.8%, hence strongly supporting the economic viability of the process in India. Moving forward, detailed investigations can also be conducted to quantify the extent of the increase in lifespan after coating as compared to untreated cookstoves, hence providing a more conclusive data on the cost and performance of the titanium-based gel.

*£1=INR93.34 (Accurate as of 10th December 2019)

4.10 *Environmental Analysis*

The concentrated sodium hydroxide solution used during the pre-treatment of the metal is considered highly toxic in nature especially when dissolved in water, hence it requires extra care when being handled during the above process. Due to its sensitivity towards moisture and potential negative health effects, responsible chemical waste management is also needed when utilising titanium (IV) isopropoxide as a precursor during

the gel synthesis phase. In contrast, both citric acid and ethylene glycol used in the synthesis of the titanium-based gel coating do not require any special waste management process as it readily biodegrades in the environment, posing insignificant threat to the environment. The resulting titanium-based sol gel from the reaction between the three reagents is identified to be environmentally friendly due to the high binding affinity and inert nature of the gel.

5. Conclusion

The Pechini method for sol-gel synthesis used in tandem with the doctor blade coating technique provides a cheap and consistent method to apply TiO_2 corrosion resistant coating on 304 stainless-steel surfaces that also has potential to be scaled up for mass production. Pre-treating the stainless-steel surface with a combination of SiC sandpaper polishing and alkali bath for surface activation provided the best wettability of the sol-gel on the surface. The drying process conducted at 100°C followed by the calcination process at 500°C with a ramp speed of 6°C/min and dwelling time of 5 minutes was considered as optimal parameters in terms of the quality of the resulting coating as well as the energy consumption and cost. Applying multiple layers of coating was also discovered to provide beneficial results in reducing the area left uncoated and increasing the density of TiO_2 on coated surface. The bending test and combustion test have shown that the coating was considerably durable when pitted against real-life usage scenarios.

Other avenues that can be considered in future research would be manipulating the pH of the sol-gel with a weak base such as ammonia to investigate how it would affect the characterisation of the coating particularly the homogeneity of TiO_2 distribution on the surface, longer or more fine-tuned combustion test and further optimisation of the drying and calcination process to reduce the amount of time taken as well as the cost. Nevertheless, both the Pechini and doctor blade coating method shows great promise in increasing the lifespan of the Prakti cookstoves without greatly increasing the cost of the stove itself.

6. Acknowledgements

We would like to express our sincere gratitude to **Dr Chris Tighe** for his guidance throughout the project as well as his insights into potential avenues for the research; **Mr Parth Shah** for his kind assistance throughout the research; **Mr Ka Ho Cheung** and **Ms Patricia Carry** for their assistance with the use of the drop shape analyser, high pressure oven, furnace and Raman spectroscopy in the Analytical Lab; **Dr Qilei Song** for his guidance on the implementation of the doctor blade coating method; **Dr Paul Fennell** for his insights and constructive feedback on the research.

7. References

Ainuddin, A. & Aziz, N., 2016. Thermal Post-treatment of TiO2 Films via Sol-gel For Enhanced Corrosion Resistance. ARPN J. Eng. Appl. Sci, Issue 11(14), pp. 8698-8703.

Anon., 2019. AIP Conference Proceedings 2068, 020013.

Ansari, F., Naderi, R. & Dehghanian, C., 2015. Improvement in the corrosion resistance of Stainless Steel 304L in Sodium Chloride Solution by a Nanoclay Incorporated Silane Coating. RSC Advances, Issue 5(1), pp. 706-716.

Antunes, R. & De Oliveira, M., 2013. Corrosion in Biomass Combustion: A Materials Selection Analysis and its Interaction With Corrosion Mechanisms and Mitigation Strategies. Corrosion Science, Issue 76, pp. 6-26.

Balamurugan, A., Kannan, S. & Rajeswari, S., n.d. Evaluation of TiO2 Coatings Obtained Using the Sol–gel technique on Surgical Grade Type 316L Stainless Steel in Simulated Body Fluid. Materials Letters, Issue 59(24-25), pp. 3138-3143.

Berni, A., Mennig, M. & Schmidt, H., 2004. Doctor Blade. Sol-gel Technologies for Glass Producers and Users, pp. 89-92.

Bloomfield, E., 2014. Gender and Livelihoods Impacts of Clean Cookstoves in South Asia. Technical Report. [Online] Available at: https://cleancookstoves.org/binarydata/RESOURCE/file/000/000/363-1.pdf.

Brady, M. P., Banta, K. & Mizia, J., 2017. Alloy Corrosion Considerations in Low-cost, Clean Biomass Cookstoves for the Developing World., s.l.: Energy for Sustainable Development.

Brinker, C. et al., 1992. Review of Sol-gel Thin Film Formation. Journal of Non-Crystalline Solids, Issue 147, pp. 424-436.

Çağlar, Yasemin & Caglar & Mujdat, 2008. Preparation and Characterization of ZnO Thin Films Deposited by Sol-gel spin Coating Method. Journal of Optoelectronics and Advanced Materials, Issue 10, pp. 2578-2583.

Ćurković, L. et al., n.d. Enhancement of Corrosion Protection of AISI 304 Stainless Steel by Nanostructured Sol–gel TiO2 Films. Corrosion Science, Issue 77, pp. 176-184.

Dimesso, L., 2018. Pechini Processes: An Alternate Approach of the Sol-Gel Method, Preparation, Properties, and Applications. Handbook of Sol-Gel Science and Technology: Processing, Characterization and Applications, pp. 1067-1088.

Dimitriev, Y., Ivanova, Y. & Iordanova, R., 2008. History of Sol-gel Science and Technology. Journal of the University of Chemical Technology and Metallurgy, Issue 43(2), pp. 181-192.

Dislich, H. & Hinz, P., 1982. History and Principles of the Sol-gel Process, and Some New Multicomponent Oxide Coatings. Journal of Non-Crystalline Solid, Issue 48(1), pp. 11-16.

Energy, D. o. t. E. a., 2019. Department of the Environment and Energy. [Online]
Available at:
http://www.environment.gov.au/protection/publications/factsheet-carbon-monoxide-co

Gelest, Inc, 2019. Sol Gel Applications | Sol Gel | Gelest, Inc. [Online]
Available at: https://www.gelest.com/applications/sol-gel-applications/

Kubiak, K., Wilson, M., Mathia, T. & Carval, P., 2011. Wettability Versus Roughness of Engineering Surfaces. Wear, Volume 271(3-4), pp. 523-528.

Kumar, Amit & Singh, Ranjeev & Bahuguna & Gaurav, 2016. Thin Film Coating through Sol-Gel Technique. Research Journal of Chemical Sciences, Issue 6, p. 65.

Liu, J., Yang, D., Shi, F. & Cai, Y., 2003. Sol–gel Deposited TiO2 Film on NiTi Surgical Alloy for Biocompatibility Improvement. Thin Solid Films, Issue 429(1-2), pp. 225-230.

Masuda, Y. & Kato, K., 2009. Iridescent Stand-Alone TiO2 Films Crystallized from Aqueous Solutions. Journal of Nanoscience and Nanotechnology, Issue 9.

Prakti, 2019. Prakti - Smart Stoves for Smart Cooks. [Online]
Available at: http://prakti.in/

Sajjadi, S. P., 2005. Sol-gel Process and Its Application in Nanotechnology. Journal of Polymer Engineering and Technology, Issue 13, pp. 38-41.

Shen, G., Chen, Y. & Lin, C., 2005. Corrosion Protection of 316 L Stainless Steel by a TiO2 Nanoparticle Coating Prepared by Sol–gel Method. Thin Solid Films, Issue 489(1-2), pp. 130-136.

Sigma Aldrich, n.d. Sigma Aldrich. [Online]
Available at:
https://www.sigmaaldrich.com/unitedkingdom.html

Stagi, et al., 2014. Light Induced TiO2 Phase Transformation: Correlation with Luminescent Surface Defects. Physica Status Solidi (b), Issue 252.

Sunde, T., Grande, T. & Einarsrud, M., n.d. Modified Pechini Synthesis of Oxide Powders and Thin Films. Handbook of Sol-Gel Science and Technology, pp. 1-30.

Ullattil, S. & Periyat, P., 2017. Sol-Gel Synthesis of Titanium Dioxide.

Worldwide & GfK, 2014. Time Spent Cooking Per Week Among Consumers By Country. [Online]
Available at:
https://www.statista.com/statistics/420719/time-spent-cooking-per-week-among-consumers-by-country/.

Xavier, C. &. C. C. &. C. S. &. B. M. &. M. M. &. C. M. &. L. E. &. S. A., 2004. Synthesis of ZrO2-based Ceramic Pigments. Journal of Thermal Analysis and Calorimetry, pp. 461-466.

In-Depth Systematic Approach for Synthesis of Fe, N Doped Carbon Nanoparticles for ORR Catalysis

Fumagalli Romario, G. L., Wild, I. C.
Department of Chemical Engineering, Imperial College London, U.K.

Abstract:
Owing to the cost and scarcity of Pt catalysts used for hydrogen fuel cells in vehicles, replacement with non-noble metal N-doped carbon nanoparticles is sought: M-N-C nanoparticles exhibit similar electrochemical performance in both alkaline and acidic media. Particularly Fe has been found to provide the most energetically favourable active site for ORR. Optimisation of Fe-N-C doped catalyst' synthesis steps was conducted to understand the individual influence of each stage on its structure and hence catalytic activity. This study achieved a systematic approach to Fe-N doped carbon catalyst synthesis. Various Fe dopants, pyrolysis stages, Fe concentrations, and acid leaching methods were investigated. Doping the carbon nanoparticles with $FeCl_2$ after support carbonisation and leaching the sample in HCl resulted in a catalyst with a performance similar to Pt/C catalysts in alkaline media, with an onset potential of 0.92V, half-wave potential of 0.74V vs RHE and limited current density of 5.0 mA cm^{-2}. This performance can be attributed to the high degree of N incorporation on the carbon surface, leading to an increase in atomic dispersion of Fe in highly dense $Fe\text{-}N_4$ active sites. Despite this impressive performance in alkali, the results in acidic media were not as impressive due to the instability of the catalysts' structure. Yet, this report outlines a logical synthesis route for high performing Fe-N doped carbon catalysts for ORR.

Introduction

In June 2018, the Ministry of Transportation of the United Kingdom, launched the "Road to Zero"[1] strategy. The aim of the policy is to eliminate manufacturing and sale of diesel- and petrol-powered cars by 2040[2]. To fill the gap in the market the government aims to support the development of electric and hydrogen fuel cells vehicles (FCVs). This latter technology in particular is promising, but a barrier that separates FCVs from the possibility of becoming fully viable is their price. Much of the costs involved in the production of fuel cells concerns the Pt catalyst component: about 30$/gram with a 60g of loading per fuel cell[3].

In FCVs H_2 is the main fuel, which is ionised at the anode. Air is fed to the cathode where O_2 undergoes oxygen reduction (ORR) to harvest protons and electrons, originated at the anode, producing H_2O as the only by-product. This pathway follows the dissociative "four-electron transfer" mechanism in which O_2 adsorbs on the metal surface and the O-O bond breaks. The alternative ORR pathway follows an associative "two-electron transfer" mechanism in which O_2 is adsorbed and the O-O bond does not break, leading to formation of H_2O_2. It is also because of its selectivity towards "four electrons transfer" that Pt is chosen as the preferred ORR catalyst. Furthermore, its superior performance and durability compared to other metals is dictated by the optimal trade-off between activity and bonding adsorption strength for both O^* and OH^* species, and by its resistance towards corrosion and poisoning[4].

Pt is used as catalyst for both electrodes, but, due to the sluggish kinetics of the aforementioned ORR, the cathode requires a Pt loading several times higher than the anode. This is why, in order to make FCV vehicles economically viable, many institutions are focusing on finding alternative catalysts for ORR to replace Pt, an expensive and relatively scarce element, with less pricy and more viable materials that can provide equal or improved electrochemical performance and stability.

Background

In the near past, researchers have been active in looking at many different ways to decrease or replace the Pt needed in fuel cells using various ingenious methods. It has been show that it is possible to reduce Pt content via alloying with less noble metals such as Fe, Ni, Co, and Cr. Mukerjee *et al.* (1995) showed that upon alloying, a contraction of the metal lattice occurs, reducing the Pt-Pt distance and enhancing the dissociation of O_2 upon adsorption[5]. Secondly, ORR has been observed and studied using metal-free carbon frameworks as catalysts. Yeager (1986) shows how ORR catalysis can occur on carbon surface; however, the kinetics are complicated to understand due to mechanistic variations between ORR happening on various carbon planes or functional groups, as also shown by Morocos and Yeager (1970)[6,7]. Taylor *et al.* (1975) described how the catalysis mechanism on glassy carbon is made even more complicated by its dependence on pH[8]. This makes finding materials with sufficient electrocatalytic activity towards ORR at all pH challenging due to the decomposition of carbon structures in acidic media[9,10]. Despite these challenges, Xuef *et al.* (2018) have demonstrated that here exists great potential for metal-free carbon materials to be further developed and used as effective ORR catalysts[11]. Heteroatoms have been proved to enhance catalytic activity of the carbon supports in which they are embedded[10]; N, B, or S

doped carbon see their properties changed. Maldonado and Stevenson (2005), synthesised doped carbon with N resulting in increased ORR potential and electron transfer number due to the change in charge distribution. The change in charge distribution enhances the side on adsorption of oxygen which promotes ORR and selectivity to H_2O. Both characteristics lead to higher catalytic activity[12]. Metal dopants (*e.g.* Fe) show increased or similar catalytic activity and stability in time for ORR compared to commercial Pt-C, as supported by Ferrero *et al.* (2016).

In the context of Fe-N-C doped catalysts, the key concept governing their performance is single- atom catalysis (SAC). In the case of ORR, the heterogeneous catalysis happens at the metal' surface. Obviously, smaller the metal's particle, the higher the fraction of total atoms that are exposed and thus available for catalysis. Thus improved electrocatalyst activity can be reached via maximum atoms' exposure in active sites. In order to achieve this scenario, Fe and N are the optimal heteroatoms dopants for multiple reasons. The radius of N (0.74 Å) is close to that of C (0.77 Å), and its electronegativity is higher (3.04 compared to 2.55). Therefore, N atoms are easily incorporated in the carbon matrix, and deeply modify the charge distribution in the network. N-dopant is not only observed as a substitution for a carbon atom (graphitic-N) but also it can be in a five-member ring (pyrrolic-N) or juxtaposed to a vacancy (pyridinic-N)[13]. The specific capability of N to be present in various configurations differently alters the local electronic distribution leading to distinct catalytic activities. N can further change the electronic properties of the carbon firmly incorporating individual metal centres, giving rise to highly dispersed SACs.

To date, great ORR catalytic activity has been shown by transition metals M-N-C materials. Many studies highlight the fundamental role of metal centres coordinated in the N-C matrix in the step of oxygen binding to promote ORR catalysis. In this paper, the non-noble metal investigated for doping is Fe, since it has been found to provide the most energetically favourable site for ORR[14]; Fe has also been observed between the transition metals to transport electrons to adsorbed OH species at the highest rate[15].

This promising technology represents a key leap towards replacing of Pt/C based catalysts in FCVs; state-of-the-art Fe-N doped carbon materials provide electrochemical activities that are either improved or similar to commercially available Pt/C catalysts in alkaline and acidic media respectively[10]. However, catalysts' manufacturing pathways followed in literature vary greatly and are diverse. There exist inconsistencies on how each synthesis step impacts the surface structure and thus electrocatalytic performance. This baffles searches directed at the active site complexes characterisation and makes M-N-C catalyst advancement attempts a trial and error procedure rather than a systematic approach.

In this paper there is a focus on utilising an Fe-N doped carbon structure synthesised from xylose, used as a carbon framework; melamine, a N-rich molecule; and various Fe-rich compounds to test their feasibility as an ORR catalyst. We investigate the synthesis of those compounds, which have proved catalytical activity[10], implementing an in-depth systematic approach. The aim is to tune particles' structure, via comparison of catalysts resulting from different synthetic pathways, to obtain optimal electrochemical performance. We investigated the effects on the catalysts' performance by varying the nature and concentrations of Fe rich precursors, the pyrolysis conditions, and the leaching acids in order to effectively engineer particles' morphology for ORR activity improvement.

Methods

Preparation of Carbon Nanospheres

The chosen sugar to serve as the carbon framework for doping atoms was xylose. The hydrothermal carbonisation (HTC) of this five-carbon sugar yielded mono dispersed spherical particles with an average diameter of 320 nm (**Figure 1**). It can be seen via the comparison with glucose nanoparticles (**Figure S1**), that the size and degree of agglomeration varies significantly. Xylose was initially chosen as optimal due to the high surface area to volume ratio which should provide good conditions for surface atomic dispersion of Fe precursors and active site formation, as proven by Yan et al. (2015)[16].

For each of the samples, 7g of solid xylose was dissolved in 70ml DI water in glass beakers which were then placed in a sonication bath for 5 minutes and stirred for 2 hours. The next stage of the synthesis was HTC of the xylose solutions,

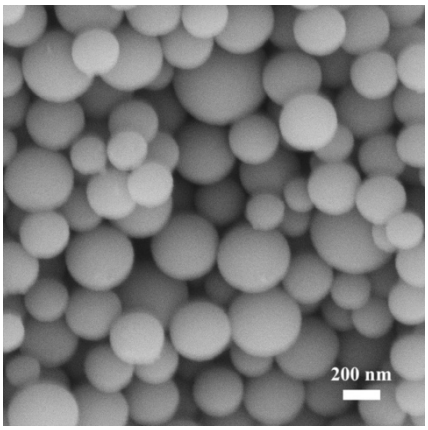


Figure 1: SEM image of Xylose-900 showing uniform particle size and low degree of agglomeration

conducted in a pressurised furnace (*Memmert, Ger*many) at 200°C for 9 hours to produce a carbon dense material. After removal from the furnace the carbonised samples still contained water, therefore separation of the hydrothermal carbon was required; this was achieved via centrifugation (*Thermo Scientific, Heraeus Multifuge X1R, USA*) for 10 minutes at 12000 rpm. After centrifugation the solids were then washed with water and placed in the centrifuge a further two times to achieve maximum separation and solid recovery. The recovered carbon was vacuum dried at 105°C for 12 hours to remove any remaining moisture, producing completely dry samples fit for milling. Milling was carried out by hand, using a pestle and mortar, to yield a fine powder of hydrothermal carbon spheres ready for doping.

Doping of Carbon Nanospheres

Five samples were synthesised using HTC spheres as a carbon support. Melamine, a N rich compound, was employed to allow the final product to have higher stability and performance[17,18]; this molecule was chosen because of the high number of N atoms present and its high availability. Fe-rich inorganic salt (*all salts are from Aldrich*) are used to introduce Fe into the structure which when coordinated to N, is believed to be a key active site for ORR[17]. Byon *et al.* (2014) found that for $FeCl_3 \bullet 6H_2O$ dopant 3wt% of nominal Fe concentration was optimal for ORR[19]. Thus, the doping of four samples to achieve 3wt% Fe in each was carried out with the addition to 500mg of HTC spheres of 500mg of melamine (*Aldrich, 99%*) and the correct amount of Fe containing dopant, that would finally result in the correct Fe percentage in the sample after STF furnace carbonisation. **Table 1** shows the dopants and the quantities used for each Fe 3wt% sample. A benchmark sample was prepared using the same quantities of HTC spheres and melamine but without an Fe-rich salt, to be used for comparison purposes. These particular Fe-rich precursors were chosen in order to analyse the effect of the nature of the dopant molecule on the catalysts structure, assuming different metal uptakes, and thus on the electrocatalytic activity. The analysis sought to understand if the optimal dopant would have been a bulkier organic salt with an already coordinated Fe-N_4 centre or a smaller molecule which could provide better surface atomic dispersion.

Dopant	Amount (mg)
Iron(II) phthalocyanine	159.38
Iron(III) phthalocyanine chloride	145.93
Iron(II) chloride	201.84
5,10,15,20-tetraphenyl-21h,23h-porphine iron(III) chloride	31.72

Table 1: summary of dopants' nature and quantity in 3wt% nanoparticles doping

First, the precursor was dissolved in 50mL EtOH and the solution bath sonicated for 10 minutes. The melamine was added, and sonication again was carried out. Finally, HTC carbon spheres were added, followed by 30 minutes of bath sonication and 2 hours of stirring. The sample without the Fe containing dopant was synthesised in the same way, adding the melamine and HTC spheres to pure EtOH. The samples were dried with the use of a rotary evaporator (*Scilogex RE100-S, USA*) first at 55°C, and 12 hours of vacuum drying at 80°C secondly. The dried samples were hand milled using a pestle and mortar to reduce particle size, and finally carbonised in a tubular furnace (*Carbolite STF, UK*) under a N_2 atmosphere at 900°C for 2 hours with a 5°C/min temperature increase. The samples synthesised via this method are named with 900 following the name of the Fe rich precursor (*e.g. $FeCl_2$-900*).

Preparation of the Working Electrodes

Inks containing the different catalysts were prepared for Rotating Disk Electrode (RDE) testing to evaluate their performance on the Oxygen Reduction Reaction (ORR). 8mg of catalyst powder was weighed and mixed with 1400µl of deionised (DI) water, 528µl of EtOH and 72µl of 5wt% Nafion©D-521 dispersion in water and 1-propanol *(Alfa Aesar)*. The Nafion© dispersion was used as binder for the components of the ink. The inks were then placed in a bath sonicator for 10 minutes before undergoing ultrasonic liquid processing in a sonic dismembrator (*Fisherbrand model 705, USA*) for 10 minutes per sample, under 5 seconds on, 5 seconds off regime. The working electrodes were then prepared by pipetting 5µl of the inks onto a polished glassy carbon electrode of diameter 3mm and left for 1 hour to dry under room temperature before electrochemical testing, resulting in a catalyst loading of 0.28mg/cm^2.

Synthesis and Doping Optimisation

After the first stage of electrochemical performance testing, the Fe precursor with the best performance towards ORR was chosen and the many methods of its synthesis were inspected and optimised in an attempt to increase the catalyst performance further.

The first step investigated the effect of acid leaching of the catalyst samples. Washing of the Fe-N carbon nanoparticles by two different methods for performance improvement was investigated. After catalyst synthesis, following the procedures outlined earlier, the first washing method consisted in the samples being leached in 1M HCl for 12 hours, followed by 1M HNO_3 for 24 hours, in order to remove respectively any iron surface impurities (*i.e.* oxides and free metals) and carbides present. Centrifugation was used for 10 minutes at 12000 rpm to separate the catalyst from the acid; the wet catalyst was then placed in the vacuum dryer at

105°C for 12 hours. The second method included only washing in 1M HCl, again for 12 hours. Based on the washing method the samples name contains 'WW' if leaching was carried out with both acids or 'W' if only one HCl washing cycle was carried out (*e.g. $FeCl_2$-[…]-W*).

The second area of optimisation involved the carbonisation technique. The finely powdered carbon nanospheres underwent a two-stage pre-doping carbonisation process under a N_2 atmosphere (600°C for 2 hours heating from room temperature at 3°C/min; then 900°C for 2 hours heating from 600°C at a rate of 2°C/min) and a further post doping carbonisation step at 900˚C for 2 hrs at 3˚C/min. These samples are named with STFX following the name of the Fe rich precursor (*e.g. $FeCl_2$-STFX*).

Finally, the effect of Fe concentration in the catalyst samples was investigated; the sample Fe wt% was changed from 3% to 1% by decreasing the mass of the Fe-rich precursor dissolved in 50mL ethanol (EtOH) *(Aldrich, 96wt%)*. Furthermore, a sample with $FeCl_2$ dopant with 0.1 Fe wt% was synthesised for comparison means during characterisation. The same methods as outlined above were then followed for the remaining synthesis. Samples synthesised via this method are named with the Fe concentration percentage following the naming code (*e.g. $FeCl_2$-[…]-0.01*).

Characterisation of Doped Materials

Characterisation of the materials included various techniques. SEM imaging was conducted using the ZEISS LEO GEMINI 1525. TEM and HAADF images were taken with FEI TITAN microscope. For XRD PANanalytical's X'PERTPRO was used. Finally, the Thermo Scientific Nexsa XPS system was employed for X-ray photoelectron spectroscopy. It should be noted that XPS data was preferentially recorded to ICP data as only surface concentrations were desired.

Testing of Doped Materials

Electrochemical tests were conducted in both acidic (50ml 0.1M $HClO_4$) and alkaline (50ml 0.1M KOH) conditions for comparison purposes in a three-electrode system, with continuous bubbling of oxygen through the electrolytes. The electrochemical data was gathered at room temperature and pressure on a rotating disk electrode (RDE) system using AUTOLAB PGSTAT 101 and a MULTI AUTOLAB M101 (CH Instruments). The configuration included three electrodes formed of: a glassy carbon 3 mm diameter RDE working electrode, a platinum rod as the counter electrode, and the saturated calomel electrode (SCE) as the reference electrode. Cyclic voltammetry (CV), linear sweep voltammetry (LSV) were completed to measure the catalyst performance towards ORR. Cyclic voltammetry was conducted in O_2 saturated solution at a scan rate of 100 mV s^{-1} in the potential sweep window of 0V to 1.2V vs RHE, linear sweep voltammetry was conducted in the same conditions, with a scan rate of 10 mV s^{-1} and a rotating electrode at 1600rpm. The potentials that were recorded in these experiments were converted to the potential vs reference hydrogen electrode (RHE) by using the following derivative of the Nernst equation:

$$E_{RHE} = E_{SCE} + 0.059pH + E^{\circ}_{SCE} \qquad (1)$$

E_{SCE} assumed as the potential vs SCE reference and E°_{SCE} = 0.245V at 20°C.

Results and Discussion

Materials Characterisation

As previously explained, the carbon nanospheres used as a support for the doping agents were synthesised from Xylose. This sugar is chosen for the high yields and degree of sphere uniformity after being exposed to hydrothermal carbonisation (HTC)[16]. After carbonisation at 900˚C, SEM images of all doped samples show similar particle size distributions (**Figure S2**), however some clear differences can be identified between the nanoparticles doped with different precursors. $FeCl_2$ seems to introduce small dispersed impurities on the nanospheres, whereas catalysts synthesised with tetraphenyl porphine Fe(II) chloride appear to have smooth homogeneous surfaces. Contrary to the previous two samples, spheres doped with Fe(II) phthalocyanine and Fe(III) phthalocyanine chloride are smaller and more agglomerated.

As a means of structure characterisation for the Fe-N doped carbon catalysts, X-ray diffraction (XRD) was completed on both doped and undoped (Xylose-900 and Xylose-STFX) samples for comparison purposes (**Figure S3**). For $FeCl_2$-900-0.03 and $FeCl_2$-STFX-0.03 the XRD plot (Figure S3) reveals a sharp peak at 2θ = 25.8° unlike the two undoped samples. This broad peak represents the (002) diffraction plane of graphite, the more intense the peak the higher the graphitic crystalline order. Hence, both of the doped samples present a higher degree of graphitisation after doping, suggesting the introduction of Fe-N doping promotes the transformation of amorphous carbon into structured crystallites during the carbonisation steps. Similarly, broad peaks at 2θ = 41° and 42° identify (100) and (101) reflection planes of less developed intergraphitic layers, highlighting the more structured nature of doped samples[2].

Another main characteristic (**Figure 2** a,b) that the dopants introduce into the carbon structure are visible surface impurities (*i.e.* oxides and free metal) as shown in SEM and TEM imaging (Figure 2 c,d,e,f). Considering the $FeCl_2$-900-0.03 sample, TEM imaging and HAADF analysis highlight a good overall degree of dispersion of Fe atoms

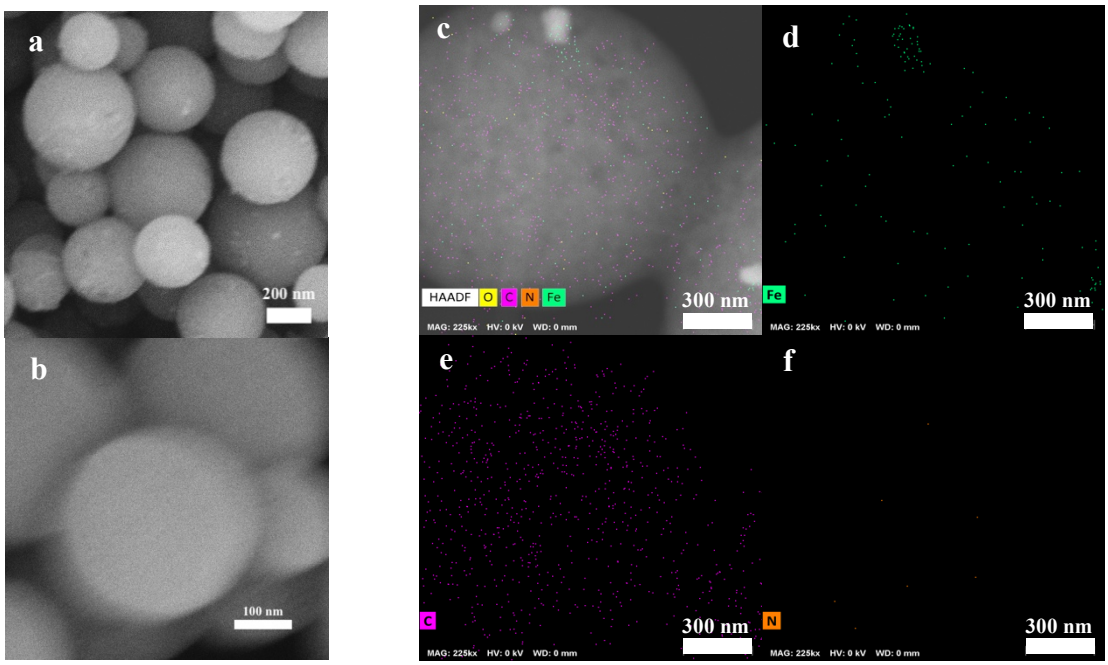


Figure 2: SEM picture of $FeCl_2$-900-0.03 sample before washing (a) showing surface impurities due to Fe particles not embedded in carbon framework and after washing (b). (c,d,e,f) HAADF imaging and TEM imaging of $FeCl_2$-900-0.03 showing good dispersion of Fe atoms on nanoparticle, highlighting their high density on sites of surface impurities.

embedded in the carbon nanosphere framework. However, it is possible to identify areas of high Fe atoms density, corresponding to aforementioned surface impurities. These impurities decrease catalytic activity modifying the surface availability and inhibiting formation of Fe-N_4 active sites[20].

Leaching the samples with HCl after the pyrolysis removes acid-soluble metal compounds that are not dispersed nor embedded in the carbon matrix, via formation of Fe chloride salts[21]. Washing with HNO_3 helps the removal of carbides as highlighted by the comparison of XRD spectra between a non-leached sample and one processed with both acids (**Figure S4**). In the leached samples. peaks at 2θ = 36° and 66°, characterising the presence of Fe_3C species[17], are not visible as in the spectra for non-leached samples. It is important to analyse the influence of the leaching step on the particles structures since there exists inconsistency in literature regarding the role of Fe_3-C sites; some state that these species inhibit the catalytic activity of single atom active sites due to their large size[15]; whereas others highlight the importance of carbides playing a key synergic role with Fe-N_x sites[22]. To find a match between the synthesis method, structure, and electrochemical performance, it was assumed that harsh leaching conditions would result in the optimal catalyst particle structure for satisfactory electrochemical activity.

Moving forward, optimisation of Fe wt% concentration was studied. Since both catalytic activity and chemoselectivity are governed by the single atom dispersion of dopants, optimising this synthesis step is crucial to achieve the sample with most uniform dispersion and thus greatest electrocatalytic activity[15]. Comparison of HAADF images can give great insight on the dispersion of Fe iron in the carbon matrix. Comparing the images for $FeCl_2$-900-0.03 (Figure 2 d) and $FeCl_2$-900-0.001 (**Figure S5**) shows that higher concentration results in the formation of impurities, where atom clusters are observed, but very low concentration results in negligible amount of Fe, and thus of active sites on the carbon surface. Hence, there must exist an optimal point at low concentration in which no formation of impurities occurs and the presence of metal atoms onto nanoparticles is ideal for single atom catalysis to happen.

Electrochemical testing of the harshly leached materials (a more detailed description in the next subsection) suggested that the HCl and HNO_3 leaching steps might have not been beneficial for the catalysts' activity. Thus, to further investigate the combined effects of concentration and leaching, XPS analysis of some samples was conducted to understand their surface chemistry (**Figure S6**a, b). Surface elemental analysis for $FeCl_2$-900-WW-0.03 and $FeCl_2$-900-WW-0.01 shows similar distinctive features for Fe and N atoms. Even considering that XPS analysis may be inaccurate at low concentrations (<0.2 atom%), the former is almost undetectable (0.1 atom% and 0.06 atom% respectively) posing a question on the optimality of the leaching system, that may be too harsh. The latter shows two clear peaks at ~398 and ~400 eV, which are assigned to pyridinic and quaternary

nitrogen atoms respectively. For the two samples N content is detected at 2.7 atom% and 3.53 atom%, resulting, also in this case particularly low.

In order to attempt to increase concentration of both Fe and N atoms in the carbon lattice, a novel method of sugar preparation was designed. This synthesis step change is driven by the fact that different carbonisation temperatures lead to different pore size distribution in the carbon material[2]. The aim was, via the stepwise pyrolysis described in the method section, to achieve highly heterogeneous surface structure of the carbon framework via first stabilising macro and mesopores at 600˚C, and then forming high density of micropores at 900˚C. This is done in order to promote better dispersion and thus embedding of dopant heteroatoms in the carbon framework. Careful analysis of XPS spectra for $FeCl_2$-STFX-WW-0.03 (Figure S6c) shows some key differences with $FeCl_2$-900-WW-0.03 spectra. Peaks related to pyridinic-N and pyrrolic-N are sharper and the N atom concentration is increased to 8 atom%. With respect to Fe, concentration is similarly low (0.05 atom%) but the peak at 712 eV, assigned to structures of Fe-N_x active sites, is sharp and easily recognizable. This is evidence of more efficient embedding of Fe atoms in N-C structures and thus formation of higher density of active sites on catalysts' surface. The optimisation of the sugar choice shows that samples undergoing pre-doping pyrolysis exhibit a structure more suitable for use for oxygen reduction reaction catalysis due to the increased N uptake and the selective insertion of metal atoms into key active sites.

From the attentive analysis inspection of XPS spectra it is possible to raise the issue about leaching optimality, already indicated by the particularly low Fe atom% in samples washed in harsh conditions. The Fe-2p XPS spectra for samples synthesised with both Xylose-900 and STFX-900, show no clear peak at 707 eV, which is assigned to Fe_3C structures[17]. As before stated, the role of iron carbides in ORR catalysis is controversial and not fully understood yet. The harsh leaching approach was decided assuming the disruptive nature of these complexes towards active sites and thus a contribution to decrease the electrocatalytic activity. HCl and HNO_3 washing shows to successfully remove carbides but, from following electrocatalytic measurements and previous comments about the low Fe content, it does not support the initial assumption. It seems that Fe_3C complexes may have a positive impact on the electrocatalytic performance of the Fe-N-C nanoparticles. As a result, the last step of procedure optimisation for the formulation of a systematic approach towards catalysts synthesis involves the replacement of harsh leaching technique with a milder washing conducted only in HCl. TEM imaging (**Figure 3**) shows uniform surface of $FeCl_2$-STFX-W-0.01, proving that the removal of impurities is efficient also in milder leaching conditions.

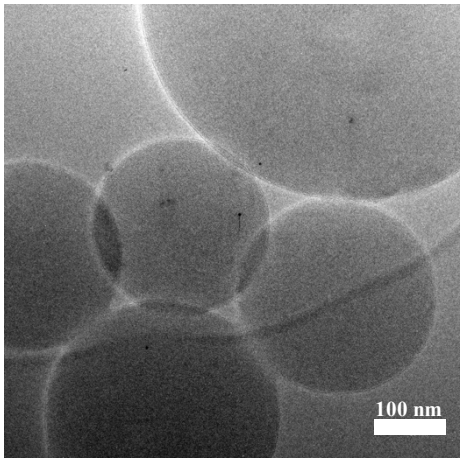


Figure 3: TEM image of $FeCl_2$-STFX-W-0.01 showing absence of surface impurities after HCl leaching

Electrocatalytic performance of doped Fe-N carbon nanospheres.

First, the investigation of the optimal synthesis pathway and preferred doping agent were conducted based on electrochemical performance in a three-electrode system for both alkaline and acidic media. The analysis of different catalytic performances was carried out initially using cyclic voltammetry (CV) in O_2 saturated solution at a scan rate of 100 mV s^{-1} in the potential sweep window of 0V to 1.2V vs RHE. Results for cyclic voltammetry in alkaline and acidic media (**Figure 4**a, b) are reported to identify an optimal dopant. In alkaline electrolyte, for all of the Fe dopants, well defined cathodic peaks can be observed at different voltages for each sample, suggesting ORR electrocatalytic activity in the environment. The cathodic peaks showed by $FeCl_2$-900 at 0.77V in 0.1 M KOH and at 0.6V in 0.1 M $HClO_4$ are, compared to some of the top performing Fe doped carbon materials found in the literature, only slightly shifted towards negative values (**Table S4**). These results are attributable to the hypothesis that square-planar Fe-N_4 centres show catalytic activity towards the ORR via a $4e^-$ pathway both in alkaline and acidic media[23]. The performance in $HClO_4$ is noticeably lower compared to the alkaline electrolyte and this can be explained by the work of *Alt et al.* who noticed stability issues of the catalyst in acidic media due to its decomposition in the acidic environment.

For further comparison between the ORR activities of the studied catalysts, rotating disk electrode (RDE at 1600 rpm) the LSV method with a scan rate of 10 mV s^{-1}, in the presence of oxygen, was employed. This analysis provided a deeper understanding of the relative electrochemical performance and the mass transfer limitations between the samples synthesised with the various

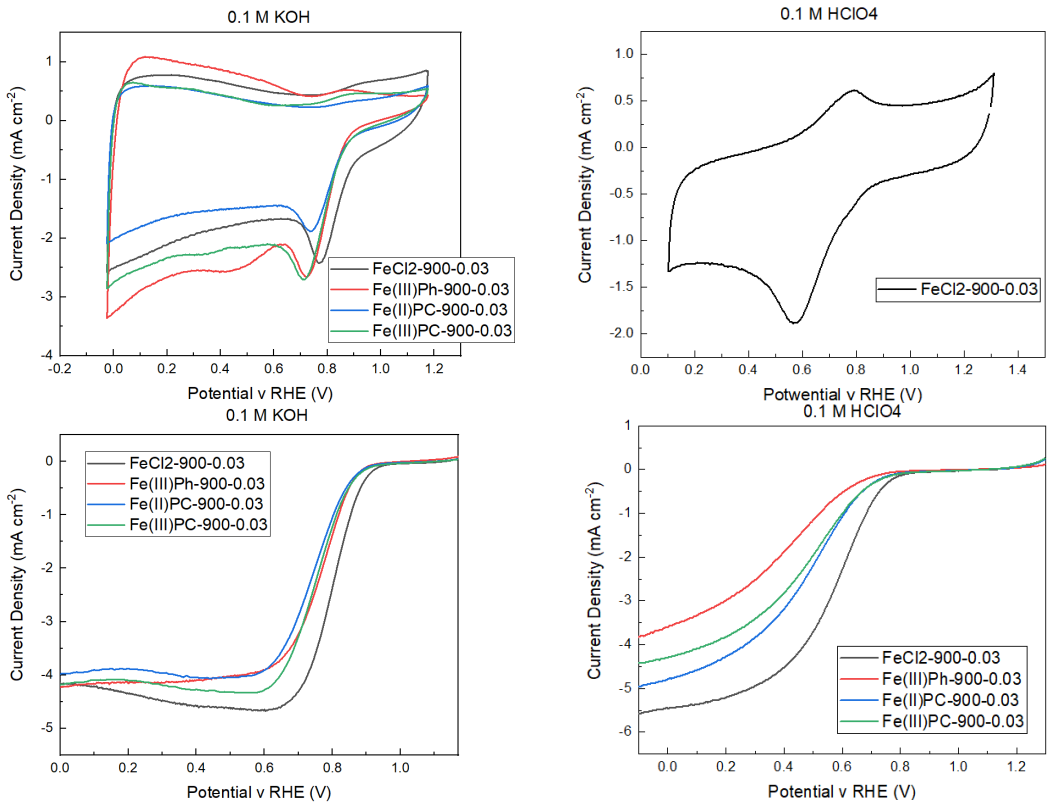


Figure 4: CV plot for all precursors in (a) 0.1M KOH and (b) 0.1M $HClO_4$. (c) LSV plot for all precursors at 1600rpm in 0.1M KOH and (d) LSV plot for $FeCl_2$ precursor at 1600rpm in 0.1M $HClO_4$.

dopants. From Figure 4c, d, onset and half-wave potentials for all Fe-N doped carbon catalysts are compared in both alkaline and acidic environments and can be found in **Table S1** and **Table S2**. Once again considering $FeCl_2$-900 as the best dopant analysed so far, in 0.1M KOH the Fe-N carbon structure shows good catalytic activity with an onset potential of 0.93V and a half-wave potential of 0.80V. When compared to state-of-art catalysts tested in literature, this performance is similar. A summary of catalytic performance of a variety of catalysts found in literature can be seen in **Table S3** and Table S4. Moreover, in acidic medium the $FeCl_2$-900 exhibits an onset potential of 0.8V and half-wave potential of 0.55V, evidently not as well performing in acidic conditions than in alkaline. Compared again to state-of-art catalysts tested in literature, this performance is not as impressive. Other interesting data that can be extracted from the LSV plot for $FeCl_2$-900 is the diffusion-limited current density which reads 4.67 mA cm^{-2} in 0.1M KOH. In 0.1M $HClO_4$ a plateau denoting O_2 diffusion-controlled region is not achieved, suggesting that the rate limiting step in the ORR is the poor reaction kinetics due to the instability of carbon material in acidic electrolytes.

$FeCl_2$ was identified as the dopant that provided the Fe-N doped carbon with the most promising electrochemical performances based on CV and LSV analysis. $FeCl_2$-900 presents more positive and similarly pronounced reduction peaks in both electrolytes and the most positive onset and halfway potential. Thus, it was chosen as the sample upon which optimisation of synthesis conditions was carried out.

Once the need for washing was established, as discussed in the previous section, and the harsh leaching conditions were assumed optimal, Fe-N doped carbon materials were washed and tested again in order to increase electrochemical performance via optimisation of Fe concentration and carbonisation techniques of the samples. Reduction of Fe concentration from 3wt% to 1wt% was aimed at providing a lower degree of metal atom agglomeration and atomic dispersion in the carbon support and thus reducing the Fe content to 1wt% was shown to slightly increase the electrochemical performance in alkaline conditions, (**Figure S7**). The $FeCl_2$-900-WW-0.01 catalyst had an onset potential of 0.9V and a half-wave potential of 0.69V, similar values to that of the $FeCl_2$-900-WW-0.03 catalyst which had an onset potential of 0.87V and a half-wave potential of 0.73V. Although both the onset and half-wave potentials were similar, it is seen that there is a substantial difference between the limited current densities, which were 4.5 mA cm^{-2} and 3.95 mA cm^{-2} of the 1wt% and 3wt% samples respectively. From this it can be said that $FeCl_2$-900-WW-0.01 has a more impressive electrochemical performance than $FeCl_2$-900-WW-0.03. This result

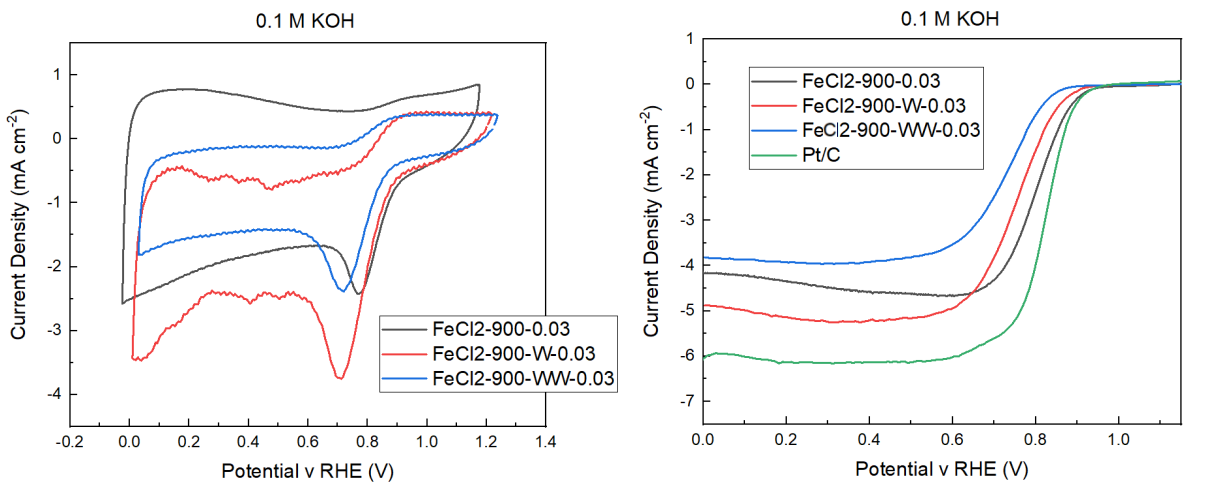


Figure 5: (a) CV plot for $FeCl_2$ precursor in 0.1M KOH after different washing techniques. (b) LSV plot for $FeCl_2$ precursor and Pt/C catalyst at 1600rpm in 0.1M KOH after different washing techniques.

implies that by decreasing the wt% of Fe, impurities were less likely to be formed on the surface as the density of Fe on atom centres available for catalysis was decreased. TEM imaging supported this and showed clusters of Fe atoms on the particles surface at higher Fe concentrations (Figure 2). This gives rise to inefficient catalysts, as not all atoms are available as reaction sites. In single atom catalysis, the accessibility of every reaction site to reactants is key for activity and selectivity[15].

After having established the synthesis pathway for the steps of doping and concentration that would result in the optimal catalyst structure, optimisation of the pyrolysis step was investigated. From the XRD plot shown in **Figure S3**, it can be seen that the different pathways of carbonisation have little to no effect on the atomic structure of the HTC spheres. However, it was assumed that the surface area and pore distribution of each of the samples would have differed, hence electrochemical tests were completed to confirm or reject this idea. **Figure S8** shows the electrochemical performance in alkaline conditions of the same $FeCl_2$ doped catalyst carbonised via the two different routes. $FeCl_2$-STFX-WW-0.01 has a greater onset potential than $FeCl_2$-900-WW-0.01; 0.94V compared to 0.9V and although the difference is not huge, having a larger onset potential is more desirable leading to a smaller overpotential and hence a quicker reaction. $FeCl_2$-STFX-WW-0.01 has a half-wave potential of 0.72V and a limited current density of 4.34V. $FeCl_2$-900-WW-0.01 has a half-wave potential of 0.69V and a limited current density of 4.5V, which is larger than that of the catalyst carbonised via the two-stage method. Although, as previously mentioned, that undergoing pre-doping pyrolysis via the STFX method increases the N uptake into the carbon structure and the selective insertion of metal atoms into key active sites, unexpectedly it seems that the performance is not improved as much as first thought. However, the lack of increased performance may be as a result of the harsh acid leaching, hence this was investigated further.

Finally, due to the unexpected results obtained from changing the pyrolysis step, deeper consideration of the effect of the acid leaching step was conducted. Rather than washing all samples harshly with both HCl and HNO_3 (WW) before testing, it was next decided to only wash with HCl (W) to view the effect on the performance. **Figure 5** shows the electrochemical performance in alkaline media of the same 3wt% $FeCl_2$-900 catalyst after different washing methods displayed as CV and LSV plots. Cyclic voltammetry results are shown in Figure 5a, and all samples show well defined cathodic peaks. $FeCl_2$-900-W-0.03 shows an extremely well pronounced cathodic peak compared to the other two samples, which promises increased catalytic performance. LSV plots for the different samples in alkaline media are shown in Figure 5b, in addition to a Pt/C catalyst plot for comparison purposes. It can be seen that although the samples have similar onset potentials, their half-wave potentials and limited current densities differ quite substantially. $FeCl_2$-900-W-0.03 appears to have the best catalytic performance out of the three Fe-N doped samples (unwashed, mildly washed (W) and harshly washed (WW)), with an onset potential of 0.92V, a half-wave potential of 0.74V vs RHE and a limited current density of 5.0 mA cm^{-2}. $FeCl_2$-900-0.03 had the second most impressive performance, while $FeCl_2$-900-WW-0.03 performed the least impressively. From these results it was assumed that harsh washing with both HCl and HNO_3 resulted in high Fe species' removal and thus in a decrease in activity.

As explained previously, washing with HCl removes surface impurities such as oxides, whereas washing with HNO_3 removed carbides. It should be noted that there is a significant increase in

performance when the HNO_3 washing step is removed; an increase in limited current density, onset potential and half-wave potential, implying that Fe_3C species are not removed. It can be suggested from these results that Fe_3C species plays a significant role in the catalysis of ORR and positively impacts the catalytic performance in alkaline media. Although the performance of $FeCl_2$-900-W-0.03 was the most impressive in these electrochemical tests, it still did not outperform the Pt/C catalyst, which has an onset potential of 0.93V, a half-wave potential of 0.83V and a limited current density of 6.1 mA cm^{-2}. Despite the good performance of $FeCl_2$-900-W-0.03 in alkaline media, in acidic media this was not reciprocated, as can be shown in **Figure S9.** $FeCl_2$-900-0.03 exhibits the highest stability in acidic conditions, as shown in Figure S9, who's onset potential and half-wave potential were written previously; followed by $FeCl_2$-900-W-0.03 and then $FeCl_2$-900-WW-0.03 who's activity does not seem to stabilise as a plateau denoting the limited current density is not reached for either catalyst.

Conclusions

In this study, through the breakdown of multiple catalyst's manufacturing stages and after individual evaluation and optimisation, insight on the independent influence of the different synthesis steps on structural characteristics and thus on the catalytic performance was achieved. The structure of Fe-N doped carbon nanoparticles for ORR catalysis have been investigated at each stage to elucidate on correlations between synthesis and activity. This was done to support future research to move away from trial and error approaches and provide procedures to correctly produce their nano-architectures and adapt their surface matrices.

The first synthesis step to optimise involved the Fe-rich precursor used to dope the HTC spheres; $FeCl_2$ provided the highest degree of incorporation and single atom dispersion into the carbon framework and hence had a higher catalytic performance that the other three tested precursors. Changes to the $FeCl_2$ precursor concentration were carried out and tested, and it was found that by decreasing the Fe wt%, a larger extent of atomic dispersion in the carbon matrix could be observed, meaning impurities such as oxides, were less likely to be formed. More active sites were available for reaction and hence an increase in catalytic performance was again seen. It should be noted that this decrease in Fe wt% was beneficial for the performance of this $FeCl_2$ precursor and for other precursors there are likely different optimal concentrations that exist. Despite expecting a change in the catalytic performance of the catalysts after the two-stage carbonisation pathway (STFX), the results gave little insight to how the pyrolysis step affected structure and performance; hypothesising that formation and stabilisation of first meso and macro, and then of micropores, enhances surface area and thus dopants adsorption. Looking ahead, more work needs to be conducted on the effect of different carbonisation methods on the catalyst surface area. Finally, the acid leaching step showed to have a significant impact on the structure and performance of the catalysts. If the acid leaching was too aggressive (WW) then this results in the removal of iron species such as Fe_3C and a decrease in catalytic activity, from this it is suggested that Fe_3C species seem likely to enhance the catalytic performance rather than inhibit it. The highest performing catalyst synthesised in this work was $FeCl_2$-900-W-0.03, with an onset potential of 0.92V, a half-wave potential of 0.74V and a limited current density of 5.3 mA cm^{-2}. To conclude, following this rational systematic approach for the synthesis of Fe-N doped carbon nanoparticles will produce high performing catalysts for ORR.

Acknowledgments

The completion of this project would have not been possible without the guidance of the Titirici Group at Imperial College London. In particular we would like to express gratitude to Mr. Jingyu Feng for his efforts in supporting us and our work.

References

1. The Road to Zero: Next steps towards cleaner road transport and delivering our Industrial Strategy, The Road to Zero: Next steps towards cleaner road transport and delivering our Industrial Strategy § (2018).

2. S. G. Badie; Y. M. Temerk; M. M. Gadelrab; I. D. Abdullah. "X-Ray Diffraction Patterns of Activated Carbons Prepared under Various Conditions." (2007) *Carbon Science 8,* 2: 95–100.

3. TIAX LLC, Platinum Availability and Economics for PEMFC Commercialization, Platinum Availability and Economics for PEMFC Commercialization § (2003).

4. S. Sheng, X. Wang, X. Zhou, Y. Su, S. Riffat, and C.-J. Liu. "A Comprehensive Review of Pt Electrocatalysts for the Oxygen Reduction Reaction: Nanostructure, Activity, Mechanism and Carbon Support in PEM Fuel Cells." (2017) *Journal of Materials Chemistry A* 5, no. 5 1808–25.

5. S. Mukerjee, S. Srinivasan, M. P. Soriaga, and J. McBreen, (1995) *J. Electrochem. Soc.*, 142, 1409

6. E. Yeager. (1986) *J. of Molecular Catalysis*, 38 5-25

7. I. Morcos, E. Yeager, Kinetic studies of the oxygen-peroxide couple on pyrolytic graphite. (1970) *Electrochimica*., 15, 953-975

8. R.J. Taylor, A.A. Humffray, Electrochemical studies on glassy carbon electrodes II. Oxygen Reduction in solution of high pH (pH>10). (1975) *J. Electroanal. Chem.*, 64 ,63-94

9. Y. Yi, G. Weinberg, M. Prenzel, M. Greiner, S. Heumann, S. Becker, R. Schlögl. Electrochemical Corrosion of a glassy carbon electrode" (2017) *Catalysis Today* 295, 32-40.

10. Chen Zhongwei, Drew Higgins, Aiping Yu, Lei Zhang, and Jiujun Zhang. A Review on Non-Precious Metal Electrocatalysts for PEM Fuel Cells. (2011) *Energy & Environmental Science 4*, no. 9 : 3167–92.

11. X. Longfei, L. Yongcheng, L. Xiaofang, L. Qingtao, S. Jiaxiang, D. Huiping, D. Liming, S. Jianglan. Zigzag carbon as efficient and stable oxygen reduction electrocatalyst for proton exchange membrane fuel cells. (2018). *Nature Communications*. 9, 3819

12. S. Maldonado, K.J. Stevenson, Influence of nitrogen doping on oxygen reduction electrocatalysis at carbon nanofiber electrodes. J. (2005), *Pyhs. Chem. B.* 109 4707-4716.

13. G. N. Tomotaroh, K. Fujisawa, V. Anil, M. Terrones, and Y. T. Yeh. "Controlling Nitrogen Doping in Graphene with Atomic Precision: Synthesis and Characterization." (2019) *Nanomaterials* 9, no. 3: 425.

14. S. Ave, E. Kibena-Põldsepp, M. Borghei, and K. Tammeveski. "Electrocatalysis of Oxygen Reduction on Heteroatom-Doped Nanocarbons and Transition Metal–Nitrogen–Carbon Catalysts for Alkaline Membrane Fuel Cells." (2018) *Journal of Materials Chemistry A* 6, no. 3: 776–804.

15. W. Aiqin, J. Li, and T. Zhang. "Heterogeneous Single-Atom Catalysis." (2018) *Nature Reviews Chemistry 2*, no. 6: 65–81.

16. Y. Qiangu, and R. Li. "Synthesis and Characterization of Carbon Nanospheres Obtained by Hydrothermal Carbonization of Wood-Derived and Other Saccharides." (2015) *Trends in Renewable Energy*.

17. G.A. Ferrero, K. Preuss, A. Marinovic, A. B. Jorge, N. Mansor, D. J. L. Brett, A. B. Fuertes, M. Sevilla, M. M. Titirici, "Fe-N-Doped Carbon capsules with outstanding electrochemical performance and stability for the oxygen reduction reaction in both acid and alkaline conditions", (2016) *ACS Nano*, 10, 5922-5932.

18. R. Tjandra, W. Liu, L. Lim, A. Yu, Melamine based, n-doped carbon/reduced graphene oxide composite foam for Li-ion Hybrid Supercapacitors, (2018) *Carbon 129*, 152-158

19. G. Selvarani, N. Leonard, and S. Calabrese Barton. "Impact of Transition Metal on Nitrogen Retention and Activity of Iron–Nitrogen–Carbon Oxygen Reduction Catalysts." (2014) *Physical Chemistry Chemical Physics 16*, no. 10: 4576.

20. Y. Hui-Juan, X. Zhang, Y. Hong, H. Maleki K. Sari, Z. Zhou, S. Sun, and X. Li. "Superior Selectivity and Tolerance towards Metal-Ion Impurities of a Fe/N/C Catalyst for CO 2 Reduction." (2019) *ChemSusChem 12*, no. 17: 3988–95.

21. Osmieri, Luigi. "Transition Metal–Nitrogen–Carbon (M–N–C) Catalysts for Oxygen Reduction Reaction. Insights on Synthesis and Performance in Polymer Electrolyte Fuel Cells." (2019) *ChemEngineering 3*, no. 1: 16.

22. Z. Xiaoran, D. Lyu, Y. B. Mollamahale, F. Yu, M. Qing, S. Yin, X. Zhang, Z. Q. Tian, and P. K. Shen. "Critical Role of Iron Carbide Nanodots on 3D Graphene Based Nonprecious Metal Catalysts for Enhancing Oxygen Reduction Reaction." (2018) *Electrochimica Acta: 281* 502–9.

23. U. Tylus, Q. Jia, K. Strickland, N. Ramaswamy, A. Serov, P. Atanassov, and S. Mukerjee. "Elucidating Oxygen Reduction Active Sites in Pyrolyzed Metal–Nitrogen Coordinated Non-Precious-Metal Electrocatalyst Systems." (2014) *The Journal of Physical Chemistry C* 118, no. 17: 8999–9008.

In Situ ATR FTIR Spectroscopy of Interactions Between High-Pressure CO_2 and Polymers

Brandon Lowe, Xiao Yuan Wang

Department of Chemical Engineering, Imperial College London

December 12, 2019

Abstract

High pressure and supercritical CO_2 have attracted much interest as alternatives in polymer processing due to its minimal environmental impact and ability to affect the molecular structure and properties of polymers. In situ ATR-FTIR spectroscopy was used to study the swelling and CO_2 sorption behaviour of eight polyketones and seven polyketone blends subjected to CO_2 at various pressures. At all the CO_2 pressures studied, both the polyketones and blends exhibited a carbonyl group concentration that resulted in maximum swelling. Beyond this maximum, further increase of the carbonyl group concentration decreased the polymer swelling. The same pattern was observed when the total sorbed CO_2 was plotted as a function of carbonyl group concentration for the polyketones. It was also observed in both the total and interacting CO_2 as a function of carbonyl group concentration plots for the blends. This suggests that beyond a certain carbonyl group concentration, interchain interactions between polymer functional groups become more dominant, meaning that less sites for CO_2 sorption are available and subsequently, less swelling can take place. It is also observed that such interchain interactions are stronger in the blends than the polyketones as the decrease in CO_2 sorption beyond the maximum observed is more evident for the blends. Furthermore, the blends do not follow the linear blending model assumptions and thus indicates that they are complex non-ideal mixtures. Finally, understanding the CO_2 sorption and swelling behaviour of polymers is important to further optimize important industrial processes such as the formation of porous polymers, polymer blends and polymer impregnation.

Keywords: *Supercritical CO_2 , ATR FTIR Spectroscopy, Polyketone , Swelling , Sorption*

1 Introduction

Polymers play an integral part in our daily lives; a few of their many applications include clothing, personal care and medical products. Employing high pressure carbon dioxide (CO_2) or supercritical carbon dioxide ($scCO_2$) as an additive in polymer processing is a promising alternative to using conventional organic solvents, as carbon dioxide is readily available and can be obtained at high purity at a low cost.[1] It also incurs no solvent removal or drying costs since it is a gas at ambient conditions and is therefore easily removed from the final product.[2] Furthermore, it is safe and environmentally friendly, being non-flammable and displaying low toxicity.

Another benefit of using high pressure CO_2 or $scCO_2$ is the effect it has on the molecular structure and properties of polymers, which as a result improves polymer processing. This is due to the specific interaction of CO_2 with polymers. It has been demonstrated both experimentally and computationally that CO_2 acts as a Lewis acid in the presence of Bronsted and Lewis bases including water, amides, amines and polymers possessing electron-donating functional groups.[3] Kazarian et al.[4] observed that the bending mode (υ_2) band of CO_2 splits in the spectrum of polymers exposed to CO_2, implying that the degeneracy of the bending mode of CO_2 has been removed. This in turn is due to the carbon atom of the CO_2 molecule interacting with the electron lone pairs of the polymer functional groups. The same phenomenon was seen in studies performed by Dobrowolski and Jamróz[5] – in the infrared (IR) spectrum of CO_2 dissolved in various organic compounds that contain Lewis base functional groups including amides, amines and dimethylsulfoxide (DMSO), two υ_2 CO_2 bands were observed.

Additionally, computational studies have come to the same conclusion. Ab initio calculations performed by Nelson and Borkman[3] have shown that the specific interaction between CO_2 and the oxygen of the carbonyl group is a Lewis acid – base type interaction. Vitoux et al.[6] have also performed such calculations on models mimicking the functional groups of polymers polydimethylsiloxane (PDMS), hydroxyl-terminated polybutadiene (HTPB) and polyethylene glycol (PEG), which all contain electron donating groups. These simulations have indicated that CO_2 also interacts with the groups in a Lewis acid-base fashion.

As CO_2 can interact with Lewis bases, it can sorb into polymers containing electron donating functional groups and cause them to swell. CO_2 sorption results in increased interchain distance and chain mobility in the polymers, an effect known as plasticization, which results in the polymer behaving as though it has been heated and reduces its viscosity.[2]

This in turn significantly benefits a myriad of polymer processing operations. One example is the formation of polymer blends, where $scCO_2$ is utilised as a swelling agent to impregnate a host polymer with a monomer and initiator and subsequent polymerization takes place.[7] Using $scCO_2$, it is possible to obtain blends that are difficult to obtain using conventional methods as $scCO_2$ mimics the effect of heating polymers.[2] It also potentially allows for processing to take place at lower temperatures, which is particularly advantageous when dealing with thermally sensitive materials.[2,7]

In addition, utilising CO_2 for the preparation of porous polymers removes the need of using "porogenic" organic solvents which have a tendency to get trapped within the polymer matrix.[7] Using CO_2, it is possible to obtain products that contain no solvent residues.[7] The same is true for polymer impregnation - by using CO_2, no solvent residues remain in the final product as CO_2 can easily be removed via depressurisation.[2,8] To further optimize the usage of $scCO_2$ in the processing of polymers, it is important to have a thorough understanding of how CO_2 interacts with and affects the properties of polymers.

In this investigation, in situ ATR-FTIR spectroscopy is adopted to study polyketones with varying carbonyl group concentration and polyketone blends subjected to high pressure CO_2. Observing the changes in the spectrum of the polyketones and polyketone blends as the pressure of applied CO_2 is increased provides information on how their properties are affected as a result of being exposed to CO_2.The spectra obtained also reveal the specific interaction of CO_2 with the polyketones. Furthermore, they can be used to determine the effect of CO_2 pressure and carbonyl group concentration on CO_2 sorption and swelling of the polyketones and polyketone blends.

2 Methodology

2.1 Experimental

In order to study polyketones and blends with varying carbonyl concentration, a range of samples with different carbonyl concentrations first needed to be synthesised. Polyketone samples 1S to 8S were obtained from a previous research group. These samples were synthesised by oxidation of cis-1,4-polybutadiene rubber. The rubber was obtained from Voronezh Synthetic Rubber Plant (Voronezh, Russian Federation) and carried a SKD trademark. Oxidation reactions were performed in a high-pressure Parr reactor loaded with rubber, benzene as solvent and N_2O in excess. This excess was approximately 2.5 times the amount of C=C bonds present in the rubber. To produce a range of samples, variations in reaction time, temperature and pressure was made.

Gas chromatography, Nuclear Magnetic Resonance (NMR) and Gel Permeation Chromatography (GPC) analysis was then performed to determine polyketone properties, these results are summarised in table A.1.1. The methodology applied to produce these polyketone samples is discussed further in associated literature,[9] described above is a brief summary of the synthesis route.

To produce the polyketone 1S-XS blends, equal parts by weight (1:1) of polyketone 1S with polyketones 2S to 8S respectively were mixed together to form a total blend mass of 0.2g. Table A.1.2 summarises key estimated physical properties of the blends. Addition of 0.3 ml $CHCl_3$ (VWR, UK) was used to ensure complete mixing had occurred. Pure spectra of $CHCl_3$ were taken as a precaution to ensure solvent addition would not interfere with spectral measurements performed on the blend samples.

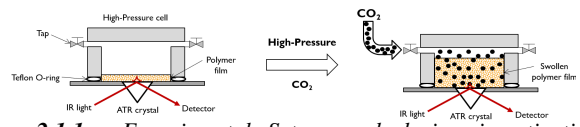


Figure 2.1.1: *Experimental Setup used during investigation, schematic reproduced with permission.*[9]

The schematic displayed in Figure 2.1.1 shows the experimental setup employed. A major advantage of spectroscopy over alternate barometric, gravimetric, visual, dynamic and thermodynamic modelling methods is that both swelling and sorption data can be measured simultaneously under the exact same experimental conditions.[10] This is not possible using any of the other prementioned methods. Moreover, ATR FTIR spectroscopic techniques provide valuable insights at the molecular level behind interactions of CO_2 and polymers[10] which may otherwise be missed by using alternate methods. Swelling and sorption can therefore be measured simultaneously and results obtained by analysing the polyketone and CO_2 anti-symmetric stretching bands respectively. ATR-FTIR spectra were obtained from a Tensor 27 FTIR spectrometer (Bruker, Germany) featuring a Mercury Cadmium Telluride single-element detector. A Golden Gate Diamond ATR accessory (Specac, UK) was utilised alongside a Teflon O-ring to create a tight seal for the high-pressure cell. The Golden Gate accessory was heated to 30°C and monitored by a temperature controller (Specac, UK). Thin films of the polyketone samples and blends were placed over the ATR diamond prior to securing of the high-pressure cell. ATR was selected over transmission mode spectroscopy since the film, although spread thinly, was too thick for transmission to effectively measure absorbance.

The high-pressure setup (HiP, USA) introduced pressurised CO_2 of 99.9% purity (BOC, UK) into the cell through a syringe pump and could be controlled by manual hand valves.

A pressure range of ambient to 50 bara was selected and measured by a E725 pressure transducer (RDP, UK). The spectral resolution was kept constant at 4 cm^{-1} with 128 coadded scans used for all obtained measurements. During data collection, three measurements were taken for each sample and blend at each pressure interval. Values for swelling and sorption for each run were then calculated and the average taken. The error bars displayed in subsequent plots use the standard deviation of the swelling and sorption results respectively.

2.2 Theory - Swelling and Sorption

2.2.1 Swelling

Swelling is defined by equation 1 as the change in volume divided by the original volume. Through application of the Beer-Lambert law (equation 2) this can be rewritten and developed into equation 3. S is the degree of swelling; ΔV is the change in volume V; c and c_0 are the concentrations of the polymer; A and A_0 are the measured absorption values at the carbonyl spectral band; ε is the molar absorptivity; $d_{e,u}$ and $d^0_{e,u}$ are the effective thicknesses for unpolarised light and 0 is used to distinguish ambient pressure conditions where no pressurised CO_2 has been added into the cell.

$$S = \frac{\Delta V}{V} = \frac{V + \Delta V}{V} - 1 = \frac{c_0}{c} - 1 \tag{1}$$

$$A = \varepsilon \cdot c \cdot d_{e,u} \tag{2}$$

$$S = \frac{A_0}{A} \cdot \frac{d_{e,u}}{d^0_{e,u}} - 1 \tag{3}$$

Since the molar absorptivity for the bending band of deionised water at 1643.5 cm^{-1} is known, as well as the concentration of DI water at 25°C, the Beer-Lambert law can be applied to compute the effective thickness of the setup by measuring the absorbance as shown in equation 4.

$$d_{e,u} = \frac{A_{1643.5}}{\varepsilon_{1643.5} \cdot c_{298k}} = \frac{A_{1643.5}}{21.8 M^{-1} cm^{-1} \cdot 55.34 M} \tag{4}$$

Assuming that this effective thickness remains approximately constant with pressure, swelling can be calculated by equation 5.

$$S \approx \frac{A_0}{A} - 1 \tag{5}$$

Integration of the carbonyl stretching spectral band υ(C=O) at 1710 cm^{-1} was used to obtain absorbance values necessary to determine the amount of swelling exhibited.

2.2.2 CO_2 Sorption

Through the Beer-Lambert Law, the total concentration of sorbed CO_2 can be calculated by using the CO_2 antisymmetric stretching band at 2335 cm^{-1}.

Equation 6 converts this concentration value to more applicable mol·g^{-1} units. Here c_w and c_V represent concentrations with respect to mass and volume units respectively and ρ is the sample density.

$$c_w = \frac{c_V \cdot (1 + S)}{\rho} \tag{6}$$

A potential issue with this calculation is that the total concentration of sorbed CO_2 is made up of two constituent effects, physical sorption as well as interacting (chemical) sorption. In order to determine the proportion of CO_2 forming chemical interactions with the polyketone carbonyl group, only the interacting concentration is desired. To account for this, subtraction of the physical sorption effect from the total amount of CO_2 sorbed was required. Sample 1S was chosen to be subtracted because it has the lowest concentration of carbonyl groups, therefore it was hypothesised to contain predominantly only physically sorbed CO_2. The result after subtraction provides an estimate for the concentration of CO_2 forming Lewis acid-base type interactions with the carbonyl groups in the remaining samples and blends.

3 Results and Discussion

3.1 Polymer Swelling

3.1.1 Evidence of Polymer Swelling

To calculate polymer swelling the carbonyl band at 1710 cm^{-1} was selected to be investigated. During swelling, the density of the polymer decreases and so consequently less infrared light is absorbed within the carbonyl band spectral region. This behaviour can clearly be observed from the spectrum shown in Figure 3.1.1 by the downwards shifting in the peak height for increasing pressure.

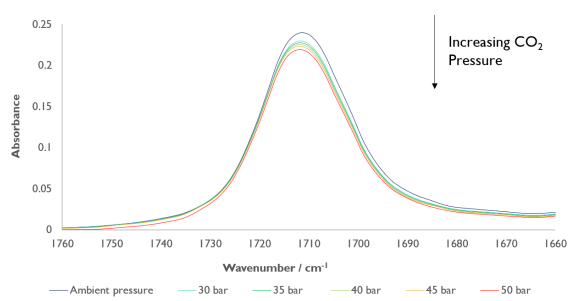


Figure 3.1.1: *ATR-FTIR spectra of polyketone 4S, showing only the carbonyl band at 1710 cm^{-1}, recorded from ambient pressure to 50 bar CO_2 at 30°C.*

3.1.2 Polyketones: Swelling

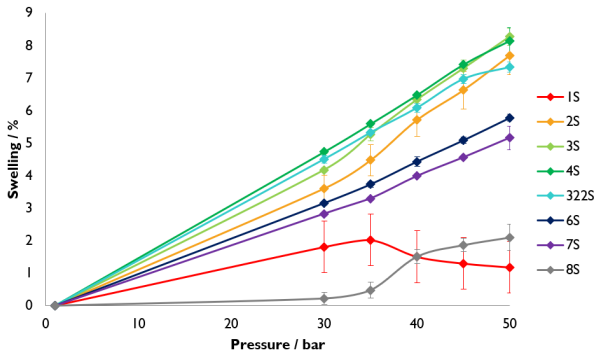


Figure 3.1.2: *Plot of polyketone swelling as a function of CO_2 pressure for polyketones 1S to 8S at 30°C.*

Figure 3.1.2 shows the swelling behaviour of the polyketone samples 1S to 8S as a function of CO_2 pressure applied. In general, the higher the pressure the greater the degree of swelling observed which is as expected. The relationship between swelling and pressure should be expected to be nonlinear and for polyketones 1S and 8S this is certainly the case, these trends agree with previous studies on these polyketones.[9] The polyketones that were observed to swell the most were samples 3S and 4S with maximum swelling values of 8.3 ± 0.3% and 8.2 ± 0.4% respectively. Plotting the same data now with respect to carbonyl group concentration produces the following trend shown in Figure 3.1.3.

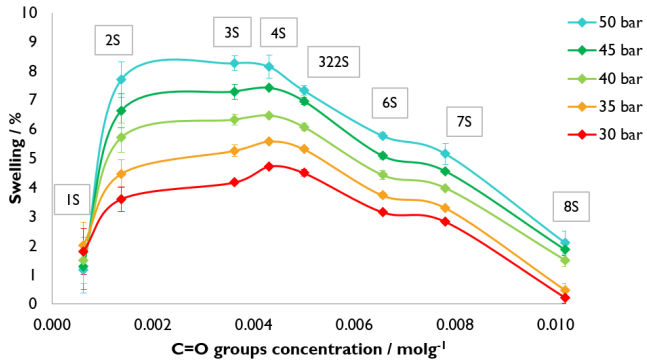


Figure 3.1.3: *Plot of polyketone swelling as a function of carbonyl group concentration for polyketones 1S to 8S; for 30 to 50 bar CO_2 pressure at 30°C.*

The previous relationship of increasing swelling with pressure can now be observed by the upwards shifting of the curve from 30 bar to 50 bar. At all pressures a maximum swelling value is obtained at around 0.004 mol·g^{-1} carbonyl concentration between polyketones 3S to 322S. Moving from 1S to 322S the increased carbonyl concentration increases the amount of CO_2-polymer interaction which subsequently leads to an increase in polymer swelling. However from 322S onwards, the carbonyl concentration becomes sufficiently high for the polymer-polymer interactions to dominate and this leads to the observed reduction in swelling. The competing effects of CO_2-polymer interactions promoting swelling and the polymer-polymer interactions preventing swelling explains the cause behind the presence of a maximum point.

3.1.3 Polyketone Blends: Swelling

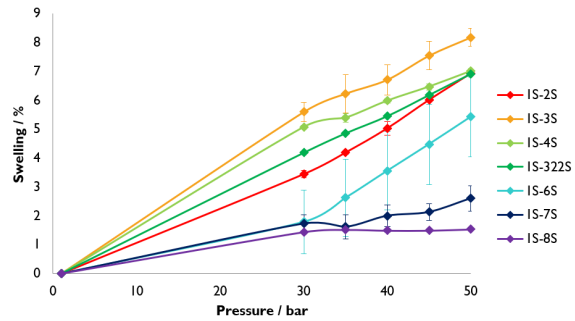


Figure 3.1.4: *Plot of polyketone blend swelling as a function of CO_2 pressure for polyketone blends 1S-XS at 30°C.*

Figure 3.1.4 displays the swelling behaviour of the polyketone 1S-XS blends with respect to pressure, note that there is a very similar trend to the polyketones 1S to 8S discussed previously in Figure 3.1.2. Again swelling is observed to increase with pressure in a non linear fashion as expected. The non linear relationship between swelling and pressure is more pronounced here for the blends. This is possibly due to the more complex nature of the interactions now between the constituent polymers (1S-1S and 1S-XS interactions) and CO_2-polymer (1S-CO_2 and XS-CO_2 interactions). Compared to the polyketone samples, the maximum amount of swelling observed reduces for the blends. This is not a surprising result, in general the addition of a comparatively low swelling polyketone like 1S to all of the other polyketone samples should be expected to reduce the degree of swelling observed. The blends with the greatest degree of swelling are, similarly to before, blends 1S-3S and 1S-4S at 8.2 ± 0.3% and 7.0 ± 0.1% respectively. For blend 1S-8S the swelling value appears to be approximately constant at 1.5% from 35 bar upwards. By observing figure 3.1.2, this is likely due to mixing the two lowest swelling polyketones together that also appear to exhibit opposite behaviour to each other. The resulting blend behaves in such a way that the competing effects of the constituent polymers effectively cancels out and an approximately constant value of swelling is hence obtained. Plotting this blend data again with respect to carbonyl group concentration produces the trends shown in Figure 3.1.5.

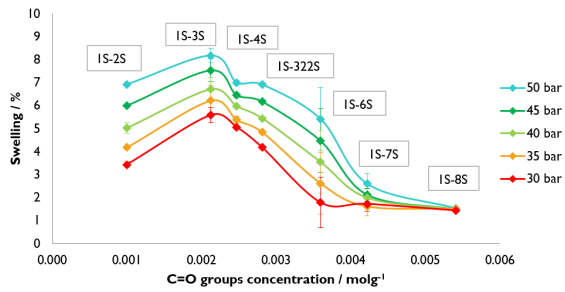


Figure 3.1.5: *Plot of polyketone blend swelling as a function of carbonyl group concentration for polyketone blends 1S-XS at 30°C.*

The point of maximum swelling is not only lower for the polyketone blends as just discussed, but also now lies at a carbonyl group concentration of approximately 0.002 mol·g^{-1} in between blends 1S-3S and 1S-4S. The explanation for the presence of the maximum point remains the same as for Figure 3.1.3. As noted previously, 1S-8S has an approximately constant value for swelling with respect to pressure. This explains why the plot for all pressure intervals converges at 1S-8S.

3.1.4 Swelling Comparison

Displayed in Figures 3.1.6 to 3.1.8 are comparisons between the constituent polyketones and the polyketone blends. The blue dashed lines represent polyketone 1S while the red dashed lines represent polyketone XS, shown in these plots are 2S, 322S and 8S respectively. If a linear blending model is assumed, then for a 1:1 blend composition, the swelling values would be expected to follow the purple dashed lines. This is halfway between the swelling values of the constituent polyketones. Such a linear blending model would only be expected for ideal mixtures. Plotted in solid green is the actual 1S-XS polyketone blends swelling behaviour. Studying these results, three important observations can be noted.

Firstly, the polyketone blends continually lie within the red and blue dashed boundaries from ambient to 50 bara. The degree of blend swelling is therefore always intermediate to the constituent polyketone swelling behaviour within this applied CO_2 pressure region. Secondly, the polyketone blend on average lies closer to one of the constituent polyketones than to the linear blending line. This result demonstrates than even for an equal mixture composition, one of the polyketones dominates the swelling behaviour of the blend. For 1S-2S and 1S-3S blends (Figures 3.1.6 and 3.1.7) the polyketone with greater carbonyl concentration dominates. However for 1S-8S (Figure 3.1.8) the reverse is observed. The results appear to suggest that, in general, whichever constituent polyketone swells the most will predominantly control the swelling behaviour of the blend. Finally, there is a clear difference between the predicted linear blending model line and the actual observed behaviour. This indicates that a non ideal mixture was produced by mixing polyketones together.

For a 1S-XS blend there would be a complex mix of 1S-1S, XS-XS, 1S-XS, 1S-CO_2 and XS-CO_2 interactions all competitively taking place. Therefore to obtain results that suggests non ideal blend mixtures were produced is to be expected.

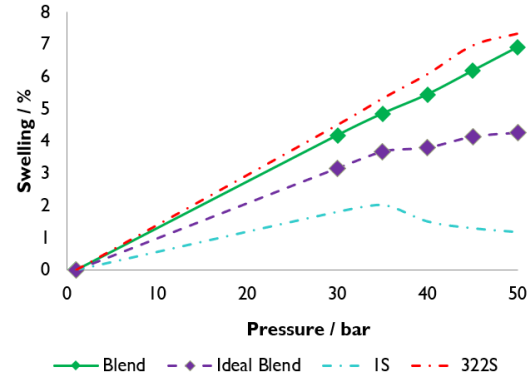


Figure 3.1.6: *Plot of polyketone swelling as a function of CO_2 pressure - Comparison of 1S, 2S and corresponding blend at 30°C.*

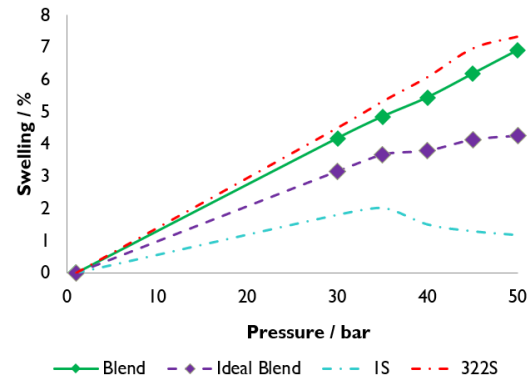


Figure 3.1.7: *Plot of polyketone swelling as a function of CO_2 pressure - Comparison of 1S, 322S and corresponding blend at 30°C.*

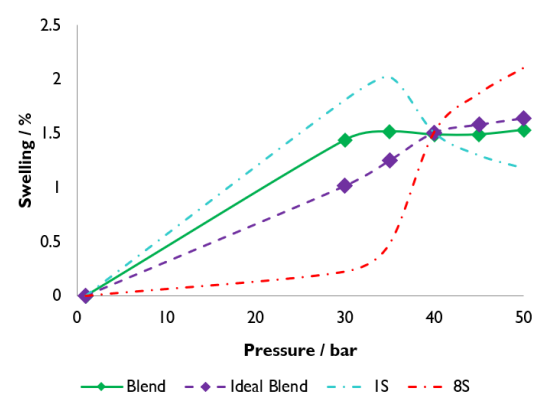


Figure 3.1.8: *Plot of polyketone swelling as a function of CO_2 pressure - Comparison of 1S, 8S and corresponding blend at 30°C.*

3.2 CO_2 Sorption

3.2.1 Changes in Spectra

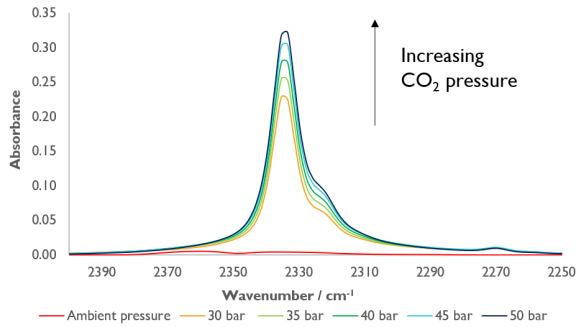


Figure 3.2.1: *ATR-FTIR spectra of the polyketone 4S, showing only the CO_2 antisymmetric (υ_3) band at 2335 cm^{-1}, recorded from ambient pressure to 50 bar CO_2 at 30°C.*

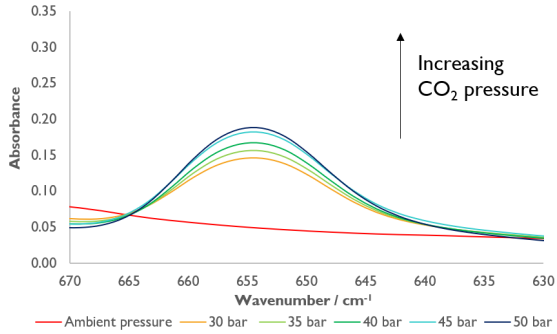


Figure 3.2.2: *ATR-FTIR spectra of the polyketone 4S, showing only the CO_2 bending (υ_2) band at 655 cm^{-1}, recorded from ambient pressure to 50 bar CO_2 at 30°C.*

Figures 3.2.1 and 3.2.2 show that, as the pressure of CO_2 applied on the polyketone 4S is increased, the absorbance of the CO_2 antisymmetric (υ_3) and bending (υ_2) bands also increase. The same pattern is observed for the rest of the polyketones and the polyketone blends. This is expected and consistent with previous studies,[9,10] as the CO_2 pressure increases, so does the concentration of CO_2 as a result of equilibrium established between the sorbed and gaseous states of CO_2.[9] The absorbance of a spectral band is proportional to the concentration according to the Beer-Lambert Law, so the higher the concentration of CO_2, the higher the absorbance of the corresponding bands.

3.2.2 Polyketones: Total CO_2 Sorption

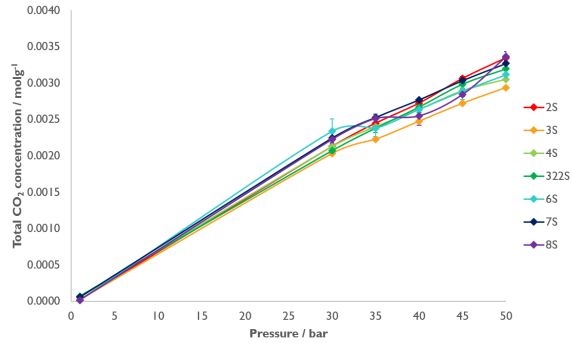


Figure 3.2.3: *Total CO_2 sorbed by the polyketones 2S to 8S as a function of CO_2 pressure at 30°C.*

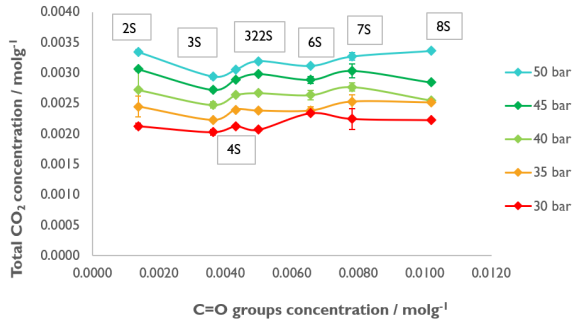


Figure 3.2.4: *Effect of polyketone carbonyl group concentration on the total CO_2 sorbed by the polyketones 2S to 8S; for 30 to 50 bar CO_2 pressure at 30°C.*

By applying the Beer-Lambert Law (equation 2) and equation 6, the concentration of total CO_2 sorbed by the polyketones can be calculated. Figure 3.2.3 demonstrates that for all polyketones, increasing the CO_2 pressure that the polyketone is exposed to results in an increase in the total CO_2 sorbed. There is little variation in total CO_2 sorption behaviour among the polyketones, as all sorb a roughly equal amount of CO_2 for all the pressures studied. Plotting the total CO_2 sorbed against polyketone carbonyl concentration (Figure 3.2.4) further emphasises this point. Increasing the carbonyl group concentration has no effect on the total CO_2 sorbed by the polyketone at all pressures. This is unlike the trends observed for polyketone swelling, where a carbonyl group concentration that resulted in maximum swelling was observed at all the pressures studied (Figure 3.1.3). To further understand such behaviour, more spectral manipulation is required, which is detailed in the following section.

3.2.3 Evidence of interacting CO_2

In the previous section, only the total sorbed concentration of CO_2 in the polyketones was calculated, which includes both physically sorbed CO_2 and CO_2 interacting with the polyketone. It is possible to perform a spectral subtraction in order to show only the bands that correspond to interacting CO_2.[9] When CO_2 interacts with the carbonyl group of the polyketone, the Lewis acid-base interaction may result in the formation of a 'bent' T-shaped carbonyl – CO_2 complex that removes the degeneracy of the CO_2 bending (υ_2) mode and results in splitting of this band in the polyketone-CO_2 spectrum.[9] As the ATR-FTIR spectrum of the polyketones exposed to CO_2 comprise of mostly bands that correspond to physically sorbed CO_2 (present at 2335, 2325 and 655 cm^{-1}), a spectral subtraction can remove these bands and reveal only the bands representing CO_2 interacting with the carbonyl groups in the polyketones.

The spectrum of 1S polyketone was chosen to be subtracted from all the polyketone spectra as it contains the most physically sorbed CO_2 due to having the least carbonyl groups.[9] This spectrum was multiplied by a variable coefficient and subtracted from the other polyketone spectrum, and two bands should be visible in the CO_2 bending mode (υ_2) region post subtraction.[11]

Figure 3.2.5 shows the 322S spectrum after subtraction of the 1S spectrum. Two bands can be seen at 662 and 650 cm^{-1}, which correspond to the out-of-plane and in-plane (nondegenerate) bending υ_2 modes respectively, thus demonstrating that there is interacting CO_2 present in the polyketone sample.[4,9]

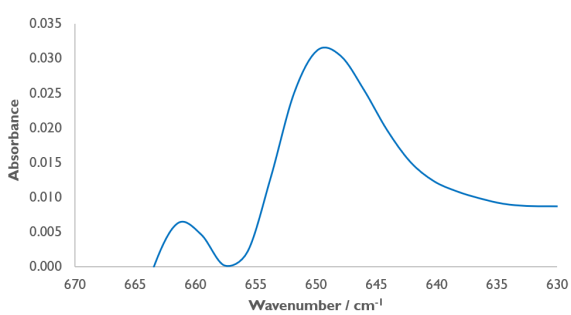


Figure 3.2.5: *Resulting spectrum upon subtraction of 1S spectrum from 322S spectrum - showing evidence of CO_2 interaction with carbonyl group in polyketone.*

3.2.4 Polyketones: Interacting CO_2

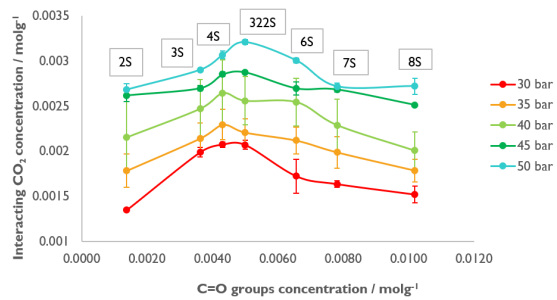


Figure 3.2.6: *Plot of interacting CO_2 for polyketones 2S to 8S as a function of carbonyl group concentration at 30°C.*

Upon subtraction of the spectrum of 1S from all the polyketone spectrum, the CO_2 antisymmetric (υ_3) band at 2335 cm^{-1} can be used to re-calculate the concentration of CO_2 via the Beer – Lambert Law. This time, it will yield the concentration of CO_2 interacting with the carbonyl group of the polyketones. Figure 3.2.6 shows the relationship between interacting CO_2 and carbonyl group concentration. Similar to what was observed in swelling (Figure 3.1.3), for all pressures there is a carbonyl group concentration that results in a maximum interacting CO_2 concentration. Furthermore, for all pressures, this maximum occurs around 0.004 mol·g^{-1} carbonyl group concentration between polyketones 2S to 322S. This again agrees with what is shown in Figure 3.1.3, and further supports the reasoning that increasing the carbonyl group concentration in the polyketones eventually leads to interactions between the carbonyl groups within the polymer chain becoming more dominant.[9] As more interchain interactions are present, this means that less carbonyl groups are available within the polymer to interact with CO_2 and subsequently, less swelling takes place.

3.2.5 Polyketone blends:Total and Interacting CO_2

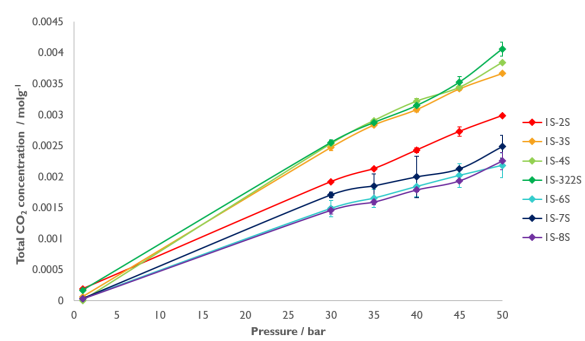


Figure 3.2.7: *Plot of total CO_2 sorbed for polyketone blends 1S-XS as a function of CO_2 pressure at 30°C.*

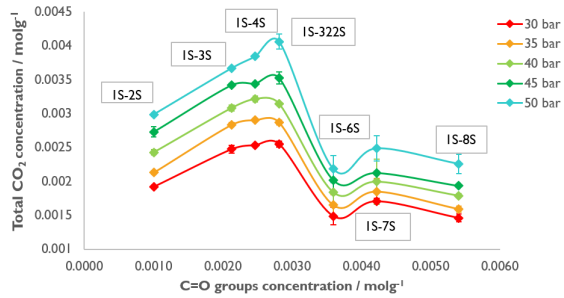


Figure 3.2.8: *Plot of total CO_2 sorbed for polyketone blends 1S-XS as a function of carbonyl group concentration at 30°C.*

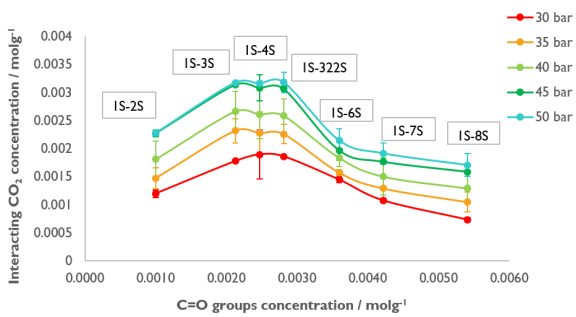


Figure 3.2.9: *Plot of interacting CO_2 for polyketone blends 1S-XS as a function of carbonyl group concentration at 30°C.*

For the polyketone blends, similar to the polyketones, increasing the CO_2 pressure results in a higher total sorbed CO_2 concentration, as shown in Figure 3.2.7. It is also observed in Figure 3.2.7 that at each pressure, there is greater variation in the total CO_2 sorbed between different blends than was seen in the single polyketones. Plotting the total CO_2 concentration as a function of blend carbonyl group concentration at different pressures (Figure 3.2.8) reveals that for all pressures, there exists a certain blend carbonyl group concentration (approximately 0.0025 mol·g^{-1}, which corresponds to blends 1S-4S to 1S-322S) that results in maximum total CO_2 being sorbed. The same phenomena is observed when the interacting CO_2 is plotted as a function of carbonyl group concentration (Figure 3.2.9) as such maxima are also present at all the pressures studied. Again, 1S-4S and 1S-322S are at the peak of the curves at all pressures, with the peak occuring at approximately 0.0025 mol·g^{-1}. This differs from the polyketones where only the interacting CO_2 plot (Figure 3.2.6) displayed maxima at all pressures. Moreover, the decrease in interacting CO_2 beyond these maxima is more significant for the blends than for the polyketones. In the blends, an additional polymer type (1S) is present. Therefore, there is greater possibility for interchain interaction in the blends – not only are there interactions between the carbonyl groups in the polymer chains but there are also additional polymer-polymer (1S-XS) interactions.

This combined effect means that there is less capacity for both total and interacting CO_2 in the blends, hence why both the total and interacting CO_2 as a function of carbonyl group concentration plots display maxima and why the decrease beyond the maxima is more prominent for the blends.

4 Conclusions and Outlook

4.1 Key Findings

The CO_2 induced swelling and sorption behaviour of eight polyketones with varying carbonyl group concentration and seven polyketone blends were studied. Increasing the applied CO_2 pressure resulted in increased swelling and total sorbed CO_2 in both the polyketones and the polyketone blends. In addition, there exists a polymer carbonyl group concentration in both the polyketones and polyketone blends (approximately 0.004 mol·g^{-1} and 0.002 mol·g^{-1} respectively) that results in maximum swelling. This maxima is observed in only the interacting CO_2 for the polyketones (around 0.004 mol·g^{-1}), but is present in both the total and interacting CO_2 for the polyketone blends (both around 0.0025 mol·g^{-1}). Additionally, decrease in swelling and sorption beyond this maxima is more evident for the blends. These observations suggest that the blends exhibit greater sensitivity than polyketones – possibly due to additional polymer - polymer interactions within the blends which lead to reduced sorption capacity. Some further comparisons should also be noted. 1S-XS blends were studied, and addition of the 1S polyketone generally reduces swelling and sorption. Additionally, when ranking in order from most to least swelling, the polyketones and their corresponding blends (for example, 3S and 1S-3S) closely match in their swelling behaviour.

$$r_s = 1 - \frac{6 \sum D^2}{n\left(n^2 - 1\right)} \tag{7}$$

Where r_s is the Spearman's Rank Correlation Coefficient

D is the difference between rankings

n is the number of pairs of ranked data points

A Spearman's rank test (equation 7) was carried out to demonstrate this and a rank coefficient of 0.94 was obtained for the swelling, showing very strong agreement. The same is true for interacting CO_2, however the agreement is not as strong as a rank coefficient of 0.76 was obtained. Finally, although polyketone blend swelling behaviour is intermediate to that of its constituent polyketones as expected, the overall blend behaviour does not follow linear blending model assumptions and indicates a complex non-ideal mixture.

4.2 Error Reduction

In this investigation, it was assumed that the effective thickness for unpolarised light ($d_{e,u}$) remained constant, as discussed in Section 2.2.1. However, as polymers swell the refractive index (RI) is likely to change. From previous studies on the polyketone samples used, a decrease in RI is observed for increasing applied CO_2 pressure.[9] The variation is approximately 0.53% within a pressure range of ambient to 50 bara.

$$\frac{d_{e,u}}{\lambda} = \left(\frac{n_2}{n_1}\cos\theta\left[3\sin^2\theta - 2\left(\frac{n_2}{n_1}\right)^2 + \left(\frac{n_2}{n_1}\right)^2\sin^2\theta\right]\right)/ \left(2\pi\left(1-\left(\frac{n_2}{n_1}\right)^2\right)\left[\left(1+\left(\frac{n_2}{n_1}\right)^2\right)\sin^2\theta - \left(\frac{n_2}{n_1}\right)^2\right]\left(\sin^2\theta - \left(\frac{n_2}{n_1}\right)^2\right)^{\frac{1}{2}}\right) \quad (8)$$

Equation 8 demonstrates that there is a link between the RI of the polyketone sample and the effective thickness. Here $d_{e,u}$ is the effective thickness for unpolarised light, λ represents the wavelength of incident light, n_1 and n_2 refer to the RI's of the ATR crystal and polymer sample respectively and θ is the angle of incidence of the light. Because this relationship exists, it is not completely accurate to assume a constant effective thickness during calculation of swelling and sorption. If two different ATR crystals are used, for example diamond and germanium, previous studies[9,10] show that a relationship can be constructed by knowing that the concentration of CO_2 is independent of the ATR crystal used. This relationship is developed from equations 9 to 11. Here d and Ge represents diamond and Germanium ATR crystals respectively. All other parameters are previously defined in section 2.2.1.

$$c_d^0 = c_{Ge}^0 \quad (9)$$

$$\frac{A_d^0}{A_{Ge}^0} = \frac{d_d^0}{d_{Ge}^0} \quad (10)$$

$$\left(\frac{A_0}{A}\cdot\frac{d_{e,u}}{d_{e,u}^0}\right)_d = \left(\frac{A_0}{A}\cdot\frac{d_{e,u}}{d_{e,u}^0}\right)_{Ge} \quad (11)$$

By taking measurements with two different ATR setups and through numerical solution of Equations 8 and 11, it is possible to compute the RI at each pressure interval. Hence instead of having to assume a constant value, the actual effective thickness with respect to pressure can be determined. This removes a source of error from the experimental results collected and would allow for more accurate swelling and sorption results to be calculated.

4.3 Future Research Potential

For industrial applications, the amount of polymer swelling or the degree of CO_2 sorption may be desired to be maximised. If so, the observed trends suggests there is an optimal carbonyl group concentration that should be aimed for. Analogous to metal alloys, polymers may be blended together in order to achieve desired physical properties and process performance targets. However, polyketone blends behave in a distinctively different manner in both swelling and sorption characteristics compared to the constituent polyketones. Furthermore, this behaviour is non ideal and does not seem to follow simple linear blending models. Predicting what polyketones should be blended together, as well as for what mixture composition, to achieve optimal swelling or sorption hence becomes a problematic question to answer. Therefore, there is a clear benefit to continuing further exploration into polyketone blends.

So far only 7 different blends at one mixture composition has been studied. Yet with the 8 polyketone samples available, it would be possible to produce 28 total blends for a given mixture composition. Two different exploratory approaches could be used for future research. Either a "breadth" based study, wherein a blend composition is fixed and all 28 polyketone blends are investigated. Or a "depth" based study, wherein different mixture compositions of the seven 1S-XS blends are made. Since the seven blends investigated so far already have a wide range of carbonyl concentrations, our personal recommendation would be to perform this latter "depth" based study.

Previous spectroscopic studies[11] have explored the effect on kinetic and thermodynamic properties due to CO_2-polymer and polymer-polymer intermolecular interactions for pure polyketone samples. Considering the chemical interaction between CO_2 and polymer functional groups, it is possible to calculate rate constants, enthalpies and entropies of such interactions and the relationship with respect to functional group concentration. However polymer blends have not yet been considered, opening up future research potential beyond swelling and sorption behaviour.

References

[1] Goodship V, Ogur EO. Polymer Processing with Supercritical Fluids. Rapra Technology Limited; 2004.

[2] Kazarian SG. Polymer Processing with Supercritical Fluids. Polymer Science. 2000;42(1):78–101.

[3] Nelson MR, Borkman RF. Ab Initio Calculations on CO_2 Binding to Carbonyl Groups. J Phys Chem A. 1998;102(40):7860–7863.

[4] Kazarian SG, Vincent MF, Bright FV, Liotta CL, Eckert CA. Specific Intermolecular Interaction of Carbon Dioxide with Polymers. J Am Chem Soc. 1996;118(7):1729–1736.

[5] Dobrowolski JC, Jamróz MH. Infrared evidence for CO_2 electron donor—acceptor complexes. Journal of Molecular Sturcture. 1992;275(1):211–219.

[6] Vitoux P, Tassaing T, Cansell F, Marre S, Aymonier C. In Situ IR Spectroscopy and Ab Initio Calculations To Study Polymer Swelling by Supercritical CO_2. J Phys Chem B. 2009;113(4):897–905.

[7] Cooper AI. Polymer synthesis and processing using supercritical carbon dioxide. Journal of Materials Chemistry. 2000;10(2):207–234.

[8] Champeau M, Thomassin JM, Jérôme C, Tassaing T. In situ FTIR micro-spectroscopy to investigate polymeric fibres under supercritical carbon dioxide: CO_2 sorption and swelling measurements. J of Supercritical Fluids. 2014;90:44–52.

[9] Ewing AV, Gabrienko AA, Semikolenov SV, Dubkov KA, Kazarian SG. How Do Intermolecular Interactions Affect Swelling of Polyketones with a Differing Number of Carbonyl Groups? An In Situ ATR-FTIR Spectroscopic Study of CO_2 Sorption in Polymers. J Phys Chem C. 2015;119(1):431–440.

[10] Flichy NMB, Kazarian SG, Lawrence CJ, Briscoe BJ. An ATR-IR Study of Poly (Dimethylsiloxane) under High-Pressure Carbon Dioxide: Simultaneous Measurement of Sorption and Swelling. J Phys Chem B. 2002;106(4):754–759.

[11] Gabrienko AA, Ewing AV, Chibiryaev AM, Agafontsev AM, Dubkov KA, Kazarian SG. New insights into the mechanism of interaction between CO_2 and polymers from thermodynamic parameters obtained by *in situ* ATR-FTIR spectroscopy. Phys Chem Chem Phys. 2016;18(9):6465–6475.

Synthesis and characterisation of hyperporous carbons from benzene-aniline copolymers for CO2 adsorption

Christopher Yeung and Fintan O'Connor
Department of Chemical Engineering, Imperial College London, U.K.

Abstract Activated carbons have been proposed as potential adsorbents for post-combustion carbon dioxide (CO_2) capture, owing to their relatively low cost and ease of preparation from a variety of materials (Stein, Wang and Fierke, 2009; Lee *et al.*, 2016). Identifying candidate sorbents with both adequate CO_2 uptake and CO_2/N_2 selectivity has been an area of significant research interest. This study investigated the performance of activated carbons derived from hypercrosslinked aniline-benzene copolymers. The produced carbons exhibited high surface area, pore volume, and CO_2 uptake. After carbonisation at 800°C, approximately 1 wt% nitrogen or less remained in the carbons. The carbon framework derived from a benzene polymer possessed the highest CO_2 uptake (4.27 mmol/g), and the greatest CO_2/N_2 selectivity (10.1) based on the IAST model. The measured CO_2 uptake is commensurate with that of zeolite 13X. The effect of the composition of the carbons on their performance could not be distinguished from other factors, as their compositions were too similar. The pore size distribution of the carbons and their ultramicropore volume (for pores less than 0.7 nm diameter), are seen to be the most significant factors determining their performance as adsorbents for CO_2.

Introduction

CO_2 accounts for approximately 76% of global greenhouse gas emissions (Centre for Climate and Energy Solutions, 2017). Carbon Capture and Storage (CCS) has been proposed as a means of reducing CO_2 emissions while satisfying growing global demands for energy and industrial output (Miller *et al.*, 2016). The development and implementation of inexpensive, effective CCS technology is critical in the medium-term future to mitigate the extent of global warming.

Currently, the most developed CCS technology is chemical absorption of CO_2 by amine scrubbing (Chang *et al.*, 2019). Despite the efficiency of the process, it has three undesirable characteristics: equipment corrosion, amine solvent degradation and associated harmful environmental waste, and a large energy penalty required for solvent regeneration. (Fayemiwo *et al.*, 2018) The energy penalty of amine scrubbing is especially problematic, as the solvent must be evaporated to remove CO_2. This penalty can reduce the electricity generation of a modern coal-fired plant by 25% (Sun *et al.*, 2015).

This highlights a need for an inexpensive, environmentally benign alternative for CCS. Using porous carbons as solid adsorbents has been put forward as a promising method. Their favourable characteristics include: high specific surface area, high thermal and chemical stability, mild adsorption condition, low manufacturing cost, ease of regeneration, low energy consumption and their environmentally benign nature (Qezelsefloo *et al.*, 2020).

Existing Methods

Research into adsorption-based CO_2 separations has focused on metal organic frameworks (MOFs), zeolites and porous aromatic frameworks (PAFs). These frameworks are synthesised with different organic ligands in order to provide a large adsorption capacity and flexible chemical structures for diverse industrial applications. For instance, MOFs have been widely applied for methane storage due to its adsorption effect and tunable properties (Sumida *et al.*, 2012); zeolite-coated ceramic membranes are used for waste-water treatment due to the high stability (Cosoli *et al.*, 2008); PAFs are studied for iodine storage with its high adsorption properties (Yan *et al.*, 2015). However, in the post-combustion process, the water composition of flue gas varies from 6.2% to 14.6% for gas and coal-fired plant (Xu *et al.*, 2003). This could pose problems to adsorbents which have low water stability such as MOFs. Water adsorption in the process will also inhibit the CO_2 adsorption of zeolites and PAFs. Considering the high synthesis costs of MOFs, PAFs and zeolites (Danaci *et al.*, 2019), they may prove difficult to scale up into CCS industrial application.

Activated carbons (ACs) have been identified as a potential sorbent for CO_2 capture, owing to their high surface area, inherent microporosity, and hydrophobicity, avoiding the need to dehumidifying flue gas stream prior to CO_2 separation (Mukherjee *et al.*, 2019). ACs may be derived inexpensively from a variety of materials, including waste materials from coffee residue to bamboo, making them a potentially accessible technology for carbon capture if proved feasible at scale (Rashidi and Yusup, 2016).

ACs may also be impregnated with heteroatoms or otherwise functionalised to improve their performance. Amine-functionalisation has been of particular interest, as such ACs may be tuneable, exploiting the CO_2 – amine interaction utilised in liquid amine scrubbing to improve CO_2 uptake on a solid sorbent (Darunte *et al.*, 2016). The presence of nitrogen in the activated carbon structure is also thought to play a role in CO_2/N_2 selectivity (Tian *et al.*, 2018).

With this goal in mind, this study derived activated carbons from nitrogen-containing,

hypercrosslinked polymers to investigate their capability as adsorbents for CO_2 capture. (Tian *et al.*, 2018).

Background

Introducing HCPs and our method

Hypercrosslinked polymers are considered in this case due to its structural flexibility, chemical stability and relatively low synthesis cost. Starting from polystyrene networks in 1970 (Tsyurupa and Davankov, 2002), HCPs have been widely recognised and studied. Through Friedel-Crafts alkylation, HCPs can be extensively linked with different precursors to produce the desired range of porosity, surface areas and functionality (Pan *et al.*, 2015). For instance, sulfonic acid-modified microporous HCPs have been studied for its potential capacity for toxic metal ions adsorption in industrial waste treatment (Li, Su, *et al.*, 2011). Therefore, suitable choices of precursors are needed for maximising the CO_2 adsorption efficiency.

Inspired by the aqueous amine-based adsorption strategy, aniline and benzene were chosen as the precursors to produce amine doped hypercrosslinked polymer to raise the CO_2 adsorption capacity and minimise the N_2 adsorption in common flue gas conditions. A 50:50 (mol%) Aniline-Benzene polymer produced by Dawson et al. has shown an outstanding CO_2/N_2 selectivity with 26:1. However, the low BET surface area of 152 m^2/g and absolute CO2 uptake of 0.85mmol/g imply that it may not have a significant impact on CO_2 adsorption on an industrial scale (Dawson *et al.*, 2012). Hence, our goal is to improve the surface area of the aniline-benzene HCPs while maintaining a satisfying CO_2/N_2 selectivity.

As aniline is the simplest aromatic amine, it was hoped that it would introduce a minimum of steric hindrance to cross-linking within the copolymers, while still introducing some nitrogen into the polymer structure. It was hoped that the derived activated carbon would possess both some nitrogen content and the high surface area and porosity of an effective adsorbent.

It is believed that carbonisation with chemical activation using an activating agent, such as KOH, would enhance the pore structure. During carbonisation, the activating agent reacts with the polymer, creating pores in the structure, improving the structural properties of the materials (Caturla, Molina-Sabio and Rodriguez-Reinoso, 1991). For example, KOH is used as an activating agent in carbonising carbon aerogels to increase the micropore volume for better hydrogen storage (Tian *et al.*, 2009). In addition, Lee et al. has shown that carbonised HCPs could have a high surface area of 4334 m^2/g (Lee *et al.*, 2016) which can possibly be adopted in our case to improve the surface areas of the aniline-benzene HCPs. Therefore, in this paper, we will introduce carbonisation to enhance the surface areas and microporosity of the aniline-benzene HCPs to improve the CO_2 adsorption capacity.

Methods

Synthesis of hypercrosslinked polymers

The copolymers were synthesised with the method described by Li et al. (Li, Gong, *et al.*, 2011), and underwent the same benzene-aniline copolymerisation as undertaken by Dawson et al. (Dawson *et al.*, 2012). Aniline and benzene were added to anhydrous 1,2-dichloroethane under a flow of nitrogen. Iron (III) chloride was added, and the reaction was heated to 80°C in an oil bath and left to continue for 23 hours. The temperature was controlled by an electronic thermometer and temperature controller. The reaction took place in a two-necked round-bottomed flask under reflux.

As in the method of Dawson et al (Dawson *et al.*, 2012), the solid product was filtered and washed with methanol until the filtrate was clear. The solid product was further washed by soxhlet extraction in methanol for 18 hours, then dried in a vacuum oven at 110°C for 24 hours.

The polymers were synthesised by Robert Woodward, a research associate in the group. One copolymer (50% benzene, 50% aniline), was synthesised by the authors.

Carbonisation of hypercrosslinked polymers

The carbonisation procedure follows Lee et al. The hypercrosslinked polymers (0.75g) were first mixed with KOH (1.50g) in a 1:2 mass ratio using pestle and mortar and the mixture was then transferred to a ceramic crucible. The samples were heated in a furnace purged with N_2 to 800°C with a 10°C/min ramp rate and was held for 2 hours. The ratio of the activating agent (Liu *et al.*, 2018) and carbonisation temperature (Lee *et al.*, 2016) were decided with the aim of maximising the surface area of the carbonised HCPs. The carbons were collected and washed with dilute HCl (~1M) and deionised water by vacuum filtration until the filtrate reached pH 7. The activated carbons were left inside vacuum oven at 110°C for a day to dry.

The polymers were labelled as B-β-A-α where β represents the mole % of benzene used during the synthesis and α represents the mole % of aniline used. For the carbons, the carbonised samples were labelled as CB-β-A-α-μ showing the precursors they were derived from. μ was assigned to carbon which was carbonised with the activating agent to polymer ratio of 4:1 in order to differentiate it from the other samples.

XPS

X-ray photoelectron spectroscopy (XPS) to determine the relative elemental composition of the carbonised HCPs, as well as the chemical states of the elements, was conducted using a Thermo Scientific K-Alpha+ X-ray photoelectron spectrometer equipped with an MXR3 Al Kα monochromated X-ray source (h_v = 1486.6 eV). The samples were initially ground and mounted onto an XPS sample holder using a small rectangular piece of conductive carbon tape. The X-ray gun power was set to 72 W (6 mA and 12 kV). Survey scans were acquired using 200 eV pass energy, 0.5 eV step size, and 100 ms (50 ms × 2 scans) dwell times. All of the high-resolution spectra (N 1s, C 1s, and O 1s, K 2p) were obtained using a 20 eV pass energy and 0.1 eV step size. The results were analysed using the Thermo Avantage data analysis program. (Shankar, Marchesini and Petit, 2019)

FTIR

The precursors were first ground to a powder using an agate mortar. Subsequently, the spectra were obtained in the wavenumber range of 500−4000 cm^{-1} using a Cary 630 FTIR Spectrometer equipped with an attenuated total reflectance accessory. (Shankar, Marchesini and Petit, 2019)

Nitrogen sorption isotherms

Nitrogen sorption isotherms were measured using a porosity and surface area analyser (Micrometrics 3 Flex) at 77K. The samples were initially degassed overnight at 120°C at approximately 0.2 mbar pressure. Subsequently, prior to the sorption isotherm measurement, the samples were degassed in situ for 5 h at 393K. The equivalent specific surface areas of the samples were determined using the Brunauer−Emmett−Teller (BET) method (Brunauer, Emmett and Teller, 1938). The total pore volume was ascertained from the volume of N_2 adsorbed at a relative pressure (P/P_0) of 0.974. The micropore volume was determined using the Dubinin−Astakhov model (Chen and Yang, 1994; Shankar, Marchesini and Petit, 2019). The above procedure was repeated for nitrogen and carbon dioxide sorption isotherms at 298K.

Thermogravimetric analysis (TGA)

The thermal stability of the samples was evaluated using a Netzch-Libra thermogravimetric analyser. In a typical analysis, the samples were loaded into an alumina crucible and subsequently heated under a nitrogen or air atmosphere (flow rate of 100 mL min^{-1}) from 30 to 900 °C with a ramp rate of 5 °C min^{-1}. (Shankar, Marchesini and Petit, 2019)

Elemental Analysis (CHNS)

Carbon, nitrogen and hydrogen contents of the precursors and carbonised HCPs were analysed by School of Human Sciences - London Metropolitan University, using a Carlo Erba Flash 2000 Elemental Analyser. Any remaining mass not ascribed to carbon, hydrogen, or nitrogen was determined to be oxygen (Petit, Seredych and Bandosz, 2009).

Results and Discussion

After the carbonisation process, the yield of the carbons ranged from 3.75% to 49.5% compared to the original polymer mass. The low yield could be subject to mass loss in the furnace due to the light weight of the products, especially the case for carbons created from precursors with higher aniline content, as well as significant mass loss from spillage during measurement and handling. However, the process would need to be repeated in order to verify the yield and also the reproducibility.

FTIR

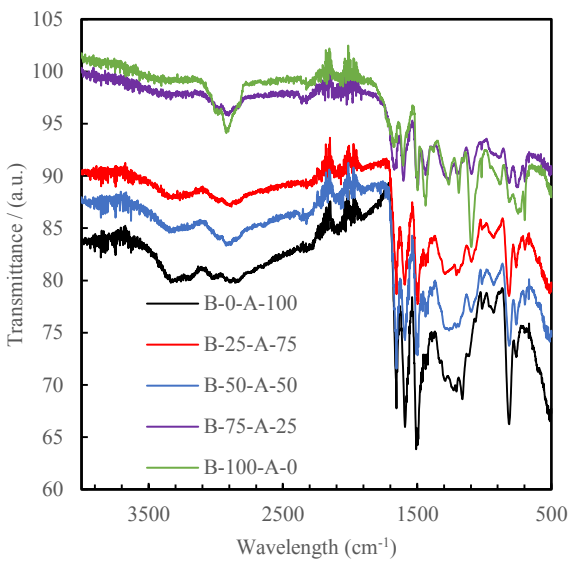
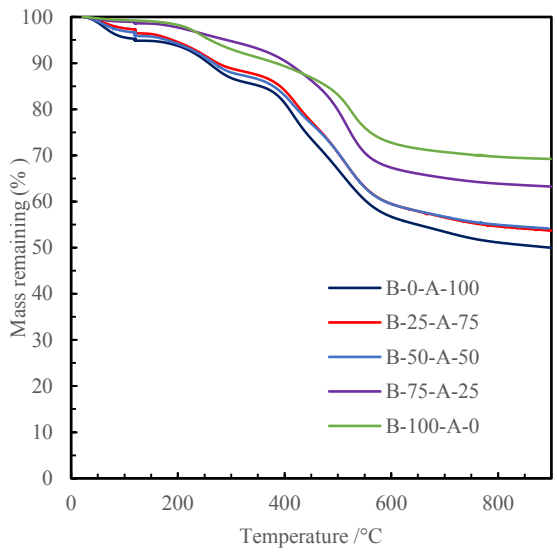


Figure 1 – FTIR spectra of hypercrosslinked polymers. Characteristic aromatic C=C bond stretching corresponds to the peaks at 1660 cm^{-1}. Primary amine N-H stretching peaks at 1580 and 3300 cm^{-1} indicate the presence of amino groups. The C-H stretching peak at 2920 cm^{-1}, corresponding to $-CH_2-$ methylene bridges, indicates the expected cross-linking.

The successful formation of hypercrosslinked polymers was confirmed by FTIR spectroscopy. The inclusion of aniline-derived amino groups in the polymers was also confirmed. The recorded spectra are shown in Figure 1. Peaks at 1660 cm^{-1} correspond to the aromatic carbon-carbon double bond stretching mode of the aromatic rings. A weak absorption peak at 3300 cm^{-1} and a strong absorption peak at 1580 cm^{-1} indicate primary amine N-H stretching. As expected, this is absent in the spectra of B-100-A-0. This shows that amino groups were

successfully translated into the polymers. Peaks at 2920 cm^{-1} correspond to C-H stretching, and were attributed to the ($-CH_2-$) methylene bridges, cross-linking the monomer units, as described in the work of Liu et al (Liu *et al.*, 2018).

A)



B)

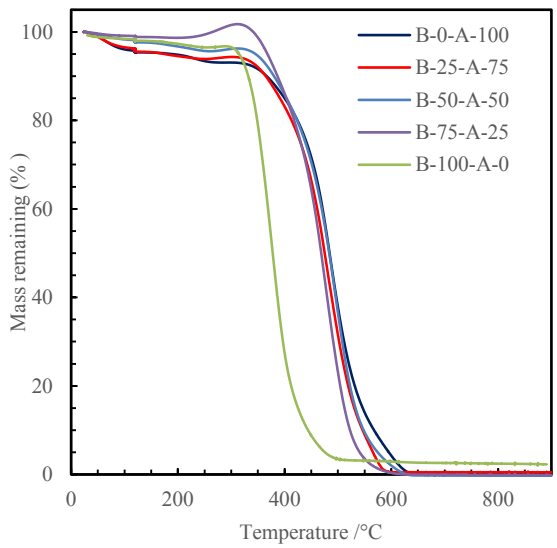


Figure 2 Thermogravimetric Analysis of HCPs:
A) Under N_2 atmosphere: The polymers become less thermally stable in direct proportion to the mole fraction of aniline used in their synthesis, suggesting weaker bonding and less cross-linking.
B) Under air: The polymers are oxidised completely by 650°C. All constituent elements are combusted or gasified, suggesting that the materials consist only of carbon, nitrogen, hydrogen and oxygen.

TGA

TGA was used to determine the thermal stability of the polymers at the temperatures used during chemical activation. Under a nitrogen atmosphere (Figure 2A), the polymers lost between 20-50% of their mass by the time the temperature reaches 800°C. The greater the proportion of aniline used to make the polymers, the lower their thermal stability under a nitrogen atmosphere. This is consistent with the expected nature of the cross-linking within the polymers. The introduction of amino groups into the polymers reduces the extent of cross-linking between monomers because of steric hindrance and the reduced availability of cross-linking sites. This makes the molecular structure of the polymer less rigid and makes it more likely that cross-linking bonds will break at high temperatures. C-B-100-A-0 made only from benzene and expected to have the greatest extent of cross-linking, is seen to be the most thermally stable. C-B-0-A-100 made only from aniline is the least thermally stable. All polymers were moderately thermally stable under nitrogen at 800°C. This shows that the polymers may undergo chemical activation without total thermal degradation.

TGA under air (Figure 2B) showed that all polymers lost 98-100% of their mass by the time the temperature reaches 800°C. This suggests that the polymers oxidise completely at high temperatures, leaving only small amounts of unburnt carbon.

Elemental Analysis

Elemental combustion analysis (CHNS) conducted on HCPs and the produced carbons was used to determine the percentage by mass of carbon, hydrogen and nitrogen in each sample. The compositions of each sample are shown in Table 1.

The nitrogen content of the polymer increased in direct proportion to the mass fraction of aniline used in their synthesis. The trend in recorded nitrogen values agrees with those reported by Dawson et al (Dawson *et al.*, 2012). This suggests that the hypercrosslinking of different proportions of aniline and benzene into polymer networks was successful.

Table 1 CHN Elemental analysis of hypercrosslinked copolymers and porous carbons. All percentages are by mass. Carbon, hydrogen, and nitrogen mass fractions determined by combustion analysis. Oxygen mass fraction taken as the percentage of mass remaining. The carbons contain between 0-1.5 wt% nitrogen, after carbonisation.

Polymers	%C	%H	%N	%O
B-100-A-0	82.7	6.4	-	10.9
B-75-A-75	80.1	5.0	2.8	12.1
B-50-A-50	73.8	5.2	5.2	15.8
B-25-A-75	71.7	5.4	7.8	15.1
B-0-A-100	68.6	5.2	10.5	15.7

Carbons	%C	%H	%N	%O
CB-100-A-0	87.6	1.8	-	10.6
CB-75-A-75	83.8	0.7	0.9	14.6
CB-50-A-50	92.5	0.4	0.6	6.5
CB-50-A-50-4	91.2	1.5	-	7.3
CB-25-A-75	86.8	0.6	0.7	11.9
CB-0-A-100	59.7	0.9	1.5	38.0

The composition of the carbons varies much less than the composition of the polymers. In every case, approximately 90% of the original nitrogen content in the polymers was removed by carbonisation. The nitrogen content of all produced carbons was in the range 0-1.5 wt%, compared to 0-10.5 wt% in the original polymers.

As amino group chemistry is the basis of current carbon capture techniques, the possibility of retaining amino group functionality in the porous carbon motivated the choice of aniline as a base for the HCPs. This suggests that it is not possible to retain primary amino groups in the carbon through the chemical activation used in this study. This is in line with the literature, which suggests that nitrogen content is greatly diminished after carbonisation above 600°C (Fan *et al.*, 2013). Amino groups are likely to be oxidised to N_2 at high temperatures during carbonisation, by similar reactions as described by Wang et al. (Wang and Kaskel, 2012) As the nitrogen content of the carbons varied by less than 1%, it was not possible to determine the effect of the nitrogen content of the carbons on their performance as adsorbents. However, it is highly unlikely that significant differences in the performance of the carbons should be attributed to differences of less than 1% in their chemical composition.

Oxygen content has been posited as having a second-order influence on the adsorption of CO_2 (Liu *et al.*, 2018). Given TGA results showing that the polymers combusted almost completely in air, the remaining wt% recorded by CHN analysis was attributed to oxygen – the only other element expected to be present. This was confirmed by XPS analysis, which reported the atomic content.

XPS

XPS measurements conducted on the samples confirmed the elemental composition determined by CHNS and provided an insight into the chemical configuration of the carbons, as well as excluding the possibility of significant impurities after carbonisation. XPS scans confirmed that no potassium was left in the carbons after carbonisation. This is important, as the presence of potassium or other elements in the carbon structure may affect its performance. It also confirms that the carbons were washed effectively after carbonisation.

XPS scans (Supporting Information) did not detect the presence of nitrogen, broadly in agreement with the result from CHNS. While XPS is a surface characterisation technique, which does not account for heterogeneity of composition within the bulk (as CHN does), it gives a useful indication of atomic content. That the XPS measurement did not detect the presence of nitrogen suggests that not enough nitrogen is present to form regular, polymeric units, as is observed in the case of oxygen. Nitrogen may be present in the form of occasional heteroatoms within the carbon structure, if at all.

Textural properties

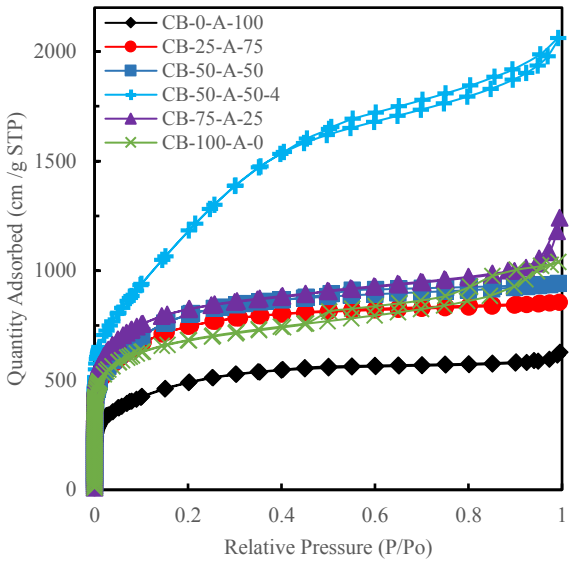


Figure 3 N_2 adsorption and desorption isotherms at 77K for all carbons.
CB-50-A-50-4, the carbon prepared using a 4:1 mass ratio of KOH:polymer, is seen to have a significantly larger uptake than those prepared with a 2:1 ratio.

The textural properties of the carbonised HCPs were summarised in Table 2. The BET surface areas and total pore volumes.of the carbons are significantly increased compared with the textural properties of the polymers reported by Dawson et al (Dawson *et al.*, 2012).This indicated that carbonisation successfully transformed the HCPs from less-porous structures to highly porous ones.

The adsorption-desorption isotherm of N_2 at 77K (Figure 3) was analysed using BET theory to determine the surface areas of the carbonised HCPs. The BET surface area of the carbons (CB-100-A-0, CB-75-A-25, CB-50-A-50, CB-25-A-75, CB-0-A-100) generally increased with the decreasing aniline content, except for CB-100-A-0. CB-75-A-25 had recorded the highest average surface area of 3074 m^2/g while CB-0-A-100 recorded the lowest of 1706 m^2/g. Surprisingly, CB-100-A-0 did not follow the trend and gave a mid-range surface area of 2530m^2/g. This result could possibly be explained by the mechanism of chemical activation of the polymers. Although aniline-derived HCPs have smaller specific surface areas than benzene-derived ones, these HCPs are less rigidly crosslinked. This provides more reaction sites for the activating agent (KOH) during carbonisation, increasing porosity and surface area.

Since the surface areas of the precursors B-0-A-100 and B-75-A-25 were already low, the surface areas of their derived carbons still increase in order of aniline used in the polymer synthesis.

Table 2 Textural properties, CO_2 and N_2 uptake and CO_2:N_2 selectivity of produced carbons. Textural properties derived from N_2 sorption at 77K.

1: BET surface area; 2: Calculated by single point pore volume at $P/P_o = 0.974$; 3:Limiting micropore volume determined by Dubinin-Astakhov method; 4: Ultramicropore volume derived from DFT pore size distribution for pores less than 0.7nm;
5: Selectivity determined by IAST: $P_{CO2} = 0.15$ bar, $P(N_2) = 0.85$ bar

Sample	Surface area [1] (m^2/g)	Pore Volume (cm^3/g)			CO_2 uptake, 1 bar [mmol/g]	N_2 uptake, 1 bar [mmol/g]	CO_2:N_2 selectivity [5]
		Total Pore Volume [2]	Micropore volume [3]	Ultramicropore volume [4]			
CB-0-A-100	1706 ±16	0.92	0.570	0.288	0.65	0.205	5.49
CB-25-A-75	2640 ± 17	1.32	0.922	0.509	2.17	0.437	4.65
CB-50-A-50	2820 ± 25	1.45	0.943	0.498	1.23	0.440	2.13
CB-50-A-50-4	4048 ± 40	3.19	1.254	0.496	2.89	0.486	7.167
CB-75-A-25	3074 ± 4.6	1.92	1.085	0.726	4.21	0.592	9.545
CB-100-A-0	2530 ±3.2	1.61	0.910	0.685	4.27	0.654	10.096

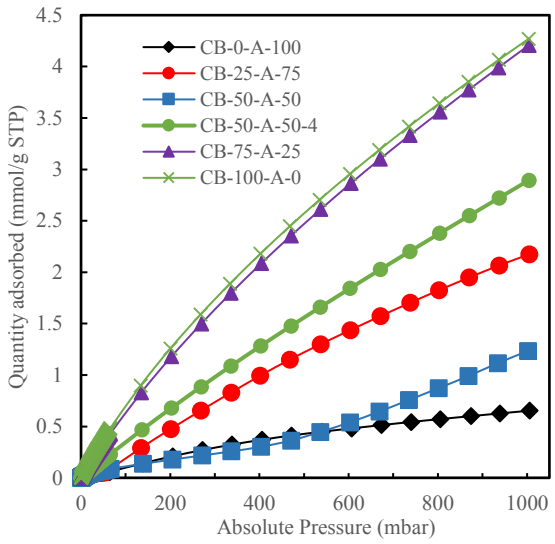


Figure 4: CO_2 adsorption isotherm at 298K for all carbons
Despite having the second-lowest surface area, and no nitrogen content, CB-100-A-0 is seen to have the greatest uptake of CO_2 (4.27 mmol/g)

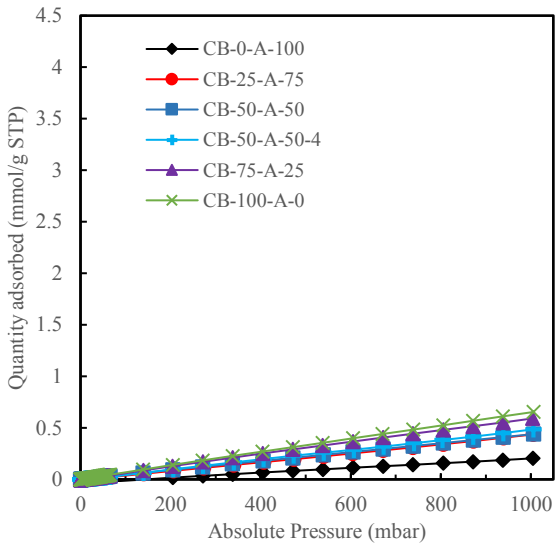


Figure 5: N_2 adsorption isotherms at 298K for all carbons
The adsorption of N_2 on the carbons is seen to be linear in pressure, showing that adsorption occurs in the Henry's Law regime.

However, this trend does not fit CB-100-A-0. The surface area of this carbon was the second lowest of the samples. This could be a result of its rigid cross-linking, which provided fewer sites for carbonisation and hence less opportunity for surface area to be enhanced. It is therefore suggested that the textural properties of the carbons depend in large part on the structure of the substance from which they are derived.

The effect of the activating agent (KOH) in carbonisation was also investigated. As shown in Table 2, the average surface area of CB-50-A-50-4 (4048 m^2/g) was 40% more than CB-50-A-50. On the other hand, the mesoporous volume for CB-50-A-50-4 (1.472 cm^3/g) was 4 times higher than CB-50-50-A (0.373 cm^3/g). Although further experiments will be required to verify the effect, the above results suggested that increasing the activating agent during carbonisation may raise both the surface area and mesoporosity of the materials.

All the carbonised HCPs showed Type I isotherms in Figure 3 which were expected and aligned with the textural properties results in Table 2. From the pore size distribution graph (Figure 6), it could be seen that most of the pore volume in the carbons except CB-50-A-50-4 were composed of pore diameters less than 2 nm, indicating that the carbonised HCPs were predominantly microporous. The total pore volume and micropore volume of the carbons followed a similar trend to the average surface area.

Adsorption Testing

The CO_2/N_2 selectivity was calculated in accordance with flue gas purification condition. As the flue gas in the industry is generally composed of 85% nitrogen and

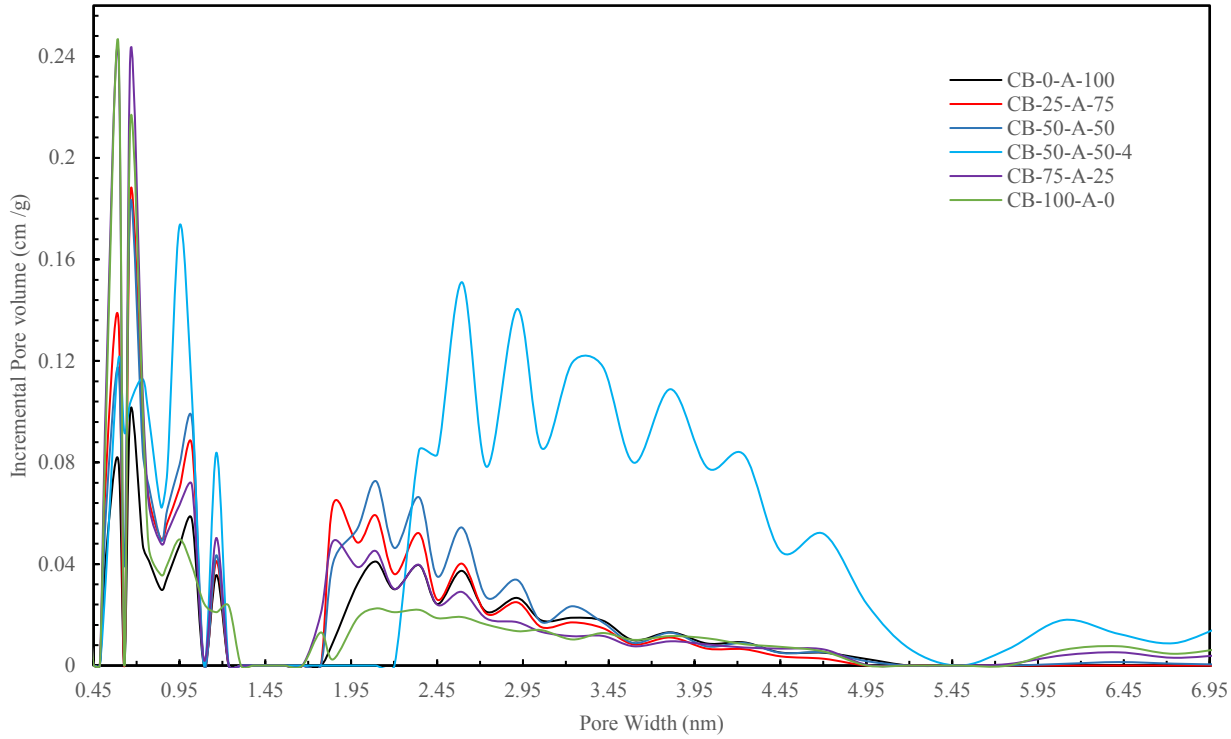


Figure 6: DFT Pore size distribution
By IUPAC definition, mesopores are 2-50 nm wide, while micropores are less than 2 nm wide. Research in activated carbons defines ultramicropores as those less than 0.7nm wide. CB-100-A-0 was seen to have the narrowest pore size distribution, with the greatest proportion of ultramicropores and the least mesoporous character. CB-75-A-25 also had significant ultramicropore volume. CB-50-A-50, which used twice as much KOH in its carbonisation as the other carbons, was seen to have significantly increased mesopore volume.

$$S = \frac{n_{CO_2}}{n_{N_2}} \cdot \frac{p_{N_2}}{p_{CO_2}} = \frac{n_{CO_2}}{n_{N_2}} \cdot \frac{0.85}{0.15}$$

Equation 1: IAST selectivity equation
Where n_i = quantity adsorbed of substance i at corresponding pressure, p_i = partial pressure of substance i (bar), S = IAST selectivity

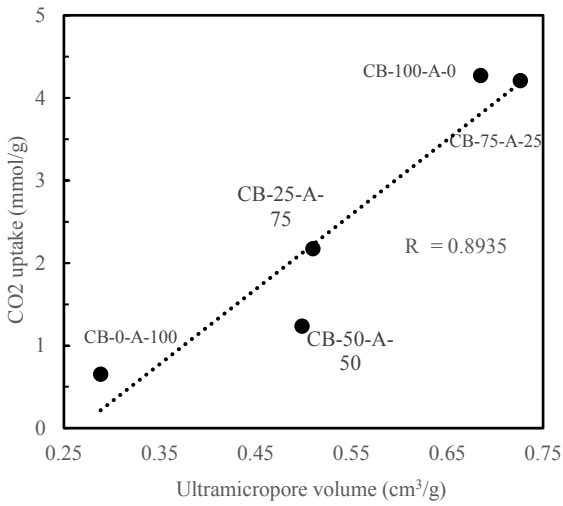


Figure 7: Correlation between CO_2 uptake and ultramicropore volume The CO_2 uptake of the carbons was seen to be strongly correlated with ultramicropore volume. This correlation is also seen in the literature (Tian *et al.*, 2018)

15% carbon dioxide (Szczęśniak and Choma, 2020), the CO_2 and N_2 adsorption capacity at the corresponding partial pressure were extracted from the isotherm (Figure 4,5) to calculate the efficient selectivity based on the IAST model (Szczęśniak and Choma, 2020).

N_2 adsorption isotherms at 298K were carried out up to 1 bar (Figure 5) to determine the adsorption capability of the carbonised HCPs. Table 2 showed the amount of N_2 adsorbed at 1 bar and the corresponding selectivity.

The absolute N_2 uptake increased steadily with decreasing aniline content. The highest N_2 uptake of 0.653 mmol/g was recorded for CB-100-A-0.

To further investigate the adsorption performance of the materials, CO_2 adsorption isotherms at 298K were carried out up to 1 bar (Figure 4). From Table 2, CB-0-A-100 had the least uptake of 0.65 mmol/g while CB-100-A-0 had the highest uptake of 4.27 mmol/g. The CO_2/N_2 selectivity of CB-100-A-0 (10.1) was also higher than CB-0-A-100 (5.49)

Although the CO_2 uptake had increased by a factor of 6 from CB-0-A-100 to CB-100-A-0, the increase in surface area was not as significant. Therefore, the pore size distribution was examined to understand the underlying causes of the increase in CO_2 uptake. From Figure 6, ones can see that CB-100-A-0 had a high proportion of micropores (< 2nm) while CB-0-A-100 had a wider distribution of pores size in the mesoporous region (2-50 nm). The difference in pore

size distribution, especially for ultramicropores (less than 0.7nm), may suggest that it is an important factor in the differences in CO_2 uptake and CO_2/N_2 selectivity.

From Figure 7, the increase in ultramicropore volume correlated strongly with increased CO_2 uptake. The large ultramicropore volume from CB-100-A-0 (0.68 cm^3/g) and CB-75-A-25 (0.73 cm^3/g) could be used to explain the high CO_2 uptake. The adsorption mechanism of ultramicropores mainly follows the micropore-filling theory. Although both CO_2 and N_2 are similar in sizes, CO_2 has a non-uniform electron charge distribution, while N_2 does not. This results in a stronger van der Waals interaction with the pore walls, enhancing the physisorption of CO_2. Therefore, CO_2 is selectively adsorbed compared to N_2. This could explain the high CO_2/N_2 selectivity and also high CO_2 uptake of CB-100-A-0 and CB-75-A-25, which both had approximately 0.2 cm^3/g more ultramicropore volume than the other carbons (Tian *et al.*, 2018).

Another factor that would potentially affect the CO_2 adsorption capacity of the carbons is the oxygen content. Although the literature reports that oxygen-containing functional groups would improve the CO_2 adsorption performance of a material(Gensterblum, Busch and Krooss, 2014; Liu *et al.*, 2018), no clear trend could be distinguished in this study. Further investigation could be carried out to identify the effect of oxygen on CO_2 adsorption for porous carbons.

Comparison of CB-100-A-0 with selected benchmarks

Lastly, the CO_2 adsorption capacity of the carbonised HCPs were compared to the current benchmark materials used in the industry. Table 3 shows the CO_2 uptake of some widely studied microporous materials for CO_2 adsorption. Although the CO_2 uptake of CB-100-A-0 was not as high as Mg-MOF-74, it has comparable results with zeolite 13X (Bezerra *et al.*, 2011; Oschatz and Antonietti, 2018), and a much higher uptake than their reported activated carbon. This result falls within our expectation as the carbonised HCPs adsorb CO_2 mainly through physisorption. Hence, they will have lower adsorption capacities than Mg-MOF-74 and UTSA-16 which undergo a chemisorption mechanism. Having similar performance to zeolite 13X suggests that the carbonised HCPs may be a good candidate for CO_2 adsorption in terms of their low production cost and benchmark comparable CO_2 adsorption capacity. However, the selectivity of CB-100-A-0 was not comparable to the industry benchmarks which suggests that our adsorbent might pose some problems when applied in the industrial process. For example, due to the relatively low selectivity, the carbons would also take up a certain amount of N_2, causing an unnecessary pressure drop. Therefore, further analysis would be needed to determine the value of our materials in the CO_2 capture industry.

Table 3 Comparison of CB-100-A-0 with selected benchmarks (Bezerra *et al.*, 2011; Oschatz and Antonietti, 2018)
CB-100-A-0 had a comparable CO_2 adsorption capacity to zeolite 13X but an unsatisfactory CO_2/N_2 selectivity when compared to other industry benchmarks (Mg-MOF-74, USTA-16, zeolite 13X)

Benchmark Sorbents	CO_2 uptake at 298K, 1 bar (mmol/g)	CO_2/N_2 selectivity at 298K, 1 bar
Mg-MOF-74	7.00	182
UTSA-16(Co)-powder	6.40	315
Zeolite 13X	4.79	500
CB-100-A-0 (Our sample)	4.27	10.1
Activated carbon	1.88	6.8

Conclusions

This study has demonstrated that the carbonisation of hypercrosslinked benzene-aniline copolymers provides a method of producing activated carbons with some properties well-suited to CO_2 adsorption. Although nitrogen content was not preserved in the activated carbons, future work could investigate the use of secondary or tertiary amines as the monomer base of the HCPs. Since the nitrogen atoms are better protected within these structures, greater nitrogen content may be preserved after carbonisation. Recent studies have investigated the use of triazine in this way (Shao *et al.*, 2018).

The carbon derived from a purely benzene-based polymer (CB-100-A-0) displayed the best CO_2 adsorption capacity and selectivity amongst the samples which implies that crosslinking benzene with aniline may not help with improving its performance. However, the textural properties obtained from the carbons suggested that fine-tuning is possible. By alternating the aniline-benzene ratio in the precursors, different textural properties can be achieved through the carbonisation.

From the results, CB-100-A-0 can potentially be a good adsorbent to be used in the CO_2 capture process as it has comparable uptake as zeolite 13X. However, further analysis needs to be carried out to determine the possibility. For instance, testing out the adsorption performance under flue gas conditions and the water uptake of the adsorbents. This will allow us to conduct a more thorough comparison with the industry benchmarks as water vapour accounts for approximately 6% in normal flue gas conditions (Xu *et al.*, 2003). An adsorbent with higher CO_2/H_2O selectivity will reduce the cost of adsorbent regeneration in the CO_2 capture process.

Previous papers have also shown that materials with ultramicropores are effective adsorbents for hydrogen and methane storage (Lozano-Castelló *et al.*, 2002; Gogotsi *et al.*, 2009). Given the high ultramicroporosity achieved in our carbon samples,

there is a potential that these materials can be used as a means of gas storage.

Lastly, the effect of temperature and activating agents can be further investigated. Studies have shown that nitrogen content would be retained with a carbonisation temperature of 600°C (Fan *et al.*, 2013). In addition, KOH would react with the nitrogen during carbonisation. Therefore, $ZnCl_2$ could be considered as the activating agent to retain the nitrogen content (Caturla, Molina-Sabio and Rodriguez-Reinoso, 1991). Implementing these changes may help with preserving the amino-CO_2 interaction to enhance CO_2/N_2 selectivity and adsorption performance.

Acknowledgement

We thank the entire Petit group enormously for welcoming us into their lab, and for their invaluable support and encouragement throughout the project.

References

Bezerra, D. P. *et al.* (2011) 'Adsorption of CO2 on nitrogen-enriched activated carbon and zeolite 13X', *Adsorption*, 17(1), pp. 235–246. doi: 10.1007/s10450-011-9320-z.

Brunauer, S., Emmett, P. H. and Teller, E. (1938) 'Adsorption of Gases in Multimolecular Layers', *Journal of the American Chemical Society*, 60(2), pp. 309–319. doi: 10.1021/ja01269a023.

Caturla, F., Molina-Sabio, M. and Rodriguez-Reinoso, F. (1991) *PREPARATION OF ACTIVATED CARBON BY CHEMICAL ACTIVATION WITH ZnCl.*

Centre for Climate and Energy Solutions (2017) *Global Emissions | Center for Climate and Energy Solutions*. Available at: https://www.c2es.org/content/international-emissions/ (Accessed: 11 December 2019).

Chang, B. *et al.* (2019) 'Poplar catkin-derived self-templated synthesis of N-doped hierarchical porous carbon microtubes for effective CO2 capture', *Chemical Engineering Journal*. Elsevier, 358, pp. 1507–1518. doi: 10.1016/J.CEJ.2018.10.142.

Chen, S. G. and Yang, R. T. (1994) 'Theoretical Basis for the Potential Theory Adsorption Isotherms. The Dubinin–Radushkevich and Dubinin–Astakhov Equations', *Langmuir*, 10(11), pp. 4244–4249. doi: 10.1021/la00023a054.

Cosoli, P. *et al.* (2008) 'Hydrogen sulphide removal from biogas by zeolite adsorption. Part I. GCMC molecular simulations', *Chemical Engineering Journal*, 145(1), pp. 86–92. doi: 10.1016/j.cej.2008.07.034.

Danaci, D. *et al.* (2019) 'Exploring the limits of adsorption-based CO_2 capture using MOFs with PVSA – from molecular design to process economics', *Molecular Systems Design & Engineering*. The Royal Society of Chemistry. doi: 10.1039/C9ME00102F.

Darunte, L. A. *et al.* (2016) 'CO2 capture via adsorption in amine-functionalized sorbents', *Current Opinion in Chemical Engineering*. Elsevier Ltd, 12, pp. 82–90. doi: 10.1016/j.coche.2016.03.002.

Dawson, R. *et al.* (2012) 'Microporous copolymers for increased gas selectivity', *Polymer Chemistry*. The Royal Society of Chemistry, 3(8), p. 2034. doi: 10.1039/c2py20136d.

Fan, X. *et al.* (2013) 'Chitosan derived nitrogen-doped microporous carbons for high performance CO2 capture', *Carbon*. Pergamon, 61, pp. 423–430. doi: 10.1016/J.CARBON.2013.05.026.

Fayemiwo, K. A. *et al.* (2018) 'Nitrogen-rich hyper-crosslinked polymers for low-pressure CO2 capture', *Chemical Engineering Journal*. Elsevier B.V., 334, pp. 2004–2013. doi: 10.1016/j.cej.2017.11.106.

Gensterblum, Y., Busch, A. and Krooss, B. M. (2014) 'Molecular concept and experimental evidence of competitive adsorption of H2O, CO2 and CH4 on organic material', *Fuel*. Elsevier Ltd, 115, pp. 581–588. doi: 10.1016/j.fuel.2013.07.014.

Gogotsi, Y. *et al.* (2009) 'Importance of pore size in high-pressure hydrogen storage by porous carbons', *International Journal of Hydrogen Energy*. Elsevier Ltd, 34(15), pp. 6314–6319. doi: 10.1016/j.ijhydene.2009.05.073.

Lee, J.-S. M. *et al.* (2016) 'Hyperporous Carbons from Hypercrosslinked Polymers', *Advanced Materials*. John Wiley & Sons, Ltd, 28(44), pp. 9804–9810. doi: 10.1002/adma.201603051.

Li, B., Gong, R., *et al.* (2011) 'A New Strategy to Microporous Polymers: Knitting Rigid Aromatic Building Blocks by External Cross-Linker', *Macromolecules*. American Chemical Society, 44(8), pp. 2410–2414. doi: 10.1021/ma200630s.

Li, B., Su, F., *et al.* (2011) 'Hypercrosslinked microporous polymer networks for effective removal of toxic metal ions from water', *Microporous and Mesoporous Materials*. Elsevier Inc., 138(1–3), pp. 207–214. doi: 10.1016/j.micromeso.2010.08.023.

Liu, M. *et al.* (2018) 'O-containing hyper-cross-linked polymers and porous carbons for CO2 capture', *Microporous and Mesoporous Materials*. Elsevier, 264, pp. 104–111. doi: 10.1016/J.MICROMESO.2017.12.036.

Lozano-Castelló, D. *et al.* (2002) 'Influence of pore size distribution on methane storage at relatively low pressure: Preparation of activated carbon with optimum pore size', *Carbon*, 40(7), pp. 989–1002. doi: 10.1016/S0008-6223(01)00235-4.

Miller, D. C. *et al.* (2016) 'Toward transformational

carbon capture systems', *AIChE Journal*. John Wiley & Sons, Ltd, 62(1), pp. 2–10. doi: 10.1002/aic.15066.

Mukherjee, A. *et al.* (2019) 'Review of post-combustion carbon dioxide capture technologies using activated carbon', *Journal of Environmental Sciences*. Elsevier, 83, pp. 46–63. doi: 10.1016/J.JES.2019.03.014.

Oschatz, M. and Antonietti, M. (2018) 'A search for selectivity to enable CO2 capture with porous adsorbents', *Energy and Environmental Science*. Royal Society of Chemistry, 11(1), pp. 57–70. doi: 10.1039/c7ee02110k.

Pan, L. *et al.* (2015) 'Polymer Chemistry one-step oxidative coupling reaction and'. Royal Society of Chemistry, pp. 2478–2487. doi: 10.1039/c4py01797h.

Petit, C., Seredych, M. and Bandosz, T. J. (2009) 'Revisiting the chemistry of graphite oxides and its effect on ammonia adsorption', *Journal of Materials Chemistry*, 19(48), pp. 9176–9185. doi: 10.1039/b916672f.

Qezelsefloo, E. *et al.* (2020) 'Adsorptive removal of CO2 on Nitrogen-doped porous carbon derived from polyaniline: Effect of chemical activation', *Materials Chemistry and Physics*. Elsevier, 239, p. 122304. doi: 10.1016/J.MATCHEMPHYS.2019.122304.

Rashidi, N. A. and Yusup, S. (2016) 'An overview of activated carbons utilization for the post-combustion carbon dioxide capture', *Journal of CO2 Utilization*. Elsevier, 13, pp. 1–16. doi: 10.1016/J.JCOU.2015.11.002.

Shankar, R., Marchesini, S. and Petit, C. (2019) 'Enhanced Hydrolytic Stability of Porous Boron Nitride via the Control of Crystallinity, Porosity, and Chemical Composition', *Journal of Physical Chemistry C*, 123(7), pp. 4282–4290. doi: 10.1021/acs.jpcc.8b11731.

Shao, L. *et al.* (2018) 'Synthesis of Triazine-Based Porous Organic Polymers Derived N-Enriched Porous Carbons for CO_2 Capture', *Industrial & Engineering Chemistry Research*. American Chemical Society, 57(8), pp. 2856–2865. doi: 10.1021/acs.iecr.7b04533.

Stein, A., Wang, Z. and Fierke, M. A. (2009) 'Functionalization of Porous Carbon Materials with Designed Pore Architecture', *Advanced Materials*. John Wiley & Sons, Ltd, 21(3), pp. 265–293. doi: 10.1002/adma.200801492.

Sumida, K. *et al.* (2012) 'Pph polkork sa', p. 8793. doi: 10.1021/cr2003272.

Sun, N. *et al.* (2015) 'Solid Adsorbents for Low-Temperature CO2 Capture with Low-Energy Penalties Leading to More Effective Integrated Solutions for Power Generation and Industrial Processes', *Frontiers in Energy Research*. Frontiers, 3, p. 9. doi: 10.3389/fenrg.2015.00009.

Szczęśniak, B. and Choma, J. (2020) 'Graphene-containing microporous composites for selective CO2 adsorption', *Microporous and Mesoporous Materials*. Elsevier, 292, p. 109761. doi: 10.1016/J.MICROMESO.2019.109761.

Tian, H. Y. *et al.* (2009) 'Enhanced hydrogen storage capacity in carbon aerogels treated with KOH', *Carbon*. Elsevier Ltd, 47(8), pp. 2128–2130. doi: 10.1016/j.carbon.2009.03.063.

Tian, Z. *et al.* (2018) 'Ultra-microporous N-doped carbon from polycondensed framework precursor for CO2 adsorption', *Microporous and Mesoporous Materials*. Elsevier, 257, pp. 19–26. doi: 10.1016/J.MICROMESO.2017.08.012.

Tsyurupa, M. P. and Davankov, V. A. (2002) 'Hypercrosslinked polymers: Basic principle of preparing the new class of polymeric materials', *Reactive and Functional Polymers*, 53(2–3), pp. 193–203. doi: 10.1016/S1381-5148(02)00173-6.

Wang, J. and Kaskel, S. (2012) 'KOH activation of carbon-based materials for energy storage', *Journal of Materials Chemistry*. The Royal Society of Chemistry, 22(45), p. 23710. doi: 10.1039/c2jm34066f.

Xu, X. *et al.* (2003) 'Separation of CO2 from Power Plant Flue Gas Using a Novel CO2 "Molecular Basket" Adsorbent', *ACS Division of Fuel Chemistry, Preprints*, 48(1), pp. 162–163.

Yan, Z. *et al.* (2015) 'Highly efficient enrichment of volatile iodine by charged porous aromatic frameworks with three sorption sites', *Angewandte Chemie - International Edition*, 54(43), pp. 12733–12737. doi: 10.1002/anie.201503362.

Prediction of equation of state data of the next-generation hydrofluoroolefin refrigerants with machine learning techniques and artificial neural networks

Alexander Baldwin & Usmaan Hafiz
Department of Chemical Engineering, Imperial College London

Abstract

In this study, an artificial neural network (ANN) model was developed to predict the temperature dependence of vapour-liquid equilibrium (VLE) properties (vapour pressure, liquid and gas densities), for a set of fourth-generation, ultra low global warming-potential hydrofluoroolefin (HFO) refrigerants. An ANN was trained using a set of pseudo-experimental data generated using the modified SAFT equation of state method set out by Mejía et al. (2014). ANN inputs were molecular structure descriptors (molecular weight and numbers of carbon and fluorine atoms), acentric factor (ω), and ab initio data computed in Gaussian 16 (dipole moment & polarisability).

The final model was capable of predicting trends in vapour pressure and density ($R^2 > 0.98$ on unseen data), but liquid density predictions were poor ($R^2 < 0.67$). It was shown that for a given HFO, an iterative method could be used with the trained ANN to estimate ω, after which a first approximation of VLE data could be found ($R^2 > 0.89$) without any experimental data. It was however demonstrated that existing correlations provide marginally better predictions than the ANN trained in this work. It is likely that ANN performance would be vastly improved if a larger experimental data set were to be used for training.

Keywords: Artificial neural networks, machine learning, hydrofluoroolefins, refrigerants, vapour-liquid equilibrium, equation of state

1 Introduction

Historically, chlorofluorocarbons (CFCs) have been widely used as refrigerants; they were first adopted as non-toxic alternatives to other haloalkanes. However, the growing scientific consensus in the 1970s on the ozone-depleting properties of CFCs prompted the phasing-out of CFCs under the 1987 Montreal Protocol (Secretariat, 2003). Research soon began into the development of suitable low ozone depletion potential (ODP) replacements. These came in the form of hydrofluorocarbons (HFCs). However, it was soon discovered that HFCs have a relatively high global warming potential (GWP) (HFC-134a, the most common refrigerant in this class had a GWP of 1,300 when compared with the impact of CO_2 over a 100-year period). A global effort to reduce the use of HFOs resulted; the Kyoto Protocol (effective 2005) listed HFCs as one of 6 greenhouse gases for which emissions should be reduced (*Kyoto Protocol to the United Nations Framework Convention on Climate Change* 1997). This was further reinforced by the 2016 Kigali Amendment to the Montreal Protocol, which introduced commitments to reduce the use and manufacture of CFCs by 85% by 2045 (*Amendment to the Montreal Protocol on Substances that Deplete the Ozone Layer* 2016).

Hydrofluoroolefins (HFOs) are of particular current interest as potential replacements for HFC refrigerants, as they have an ODP of 0 and an 'ultra-low' GWP less than 30 (Brown et al., 2016). Whilst ultra-low GWP values and an ODP of 0 are essential for new refrigerants, it is important that refrigerants have desired behaviour in thermodynamic cycles. Additionally, new refrigerants should ideally be 'backwards-compatible' with existing appliances and machinery, so that current equipment does not go to waste and existing refrigerants can be responsibly decommissioned.

Understanding thermophysical behaviour is thus critical when designing new refrigerants; during use refrigerants cycle through significant variations in temperature and pressure, affecting their state, density and enthalpy. When designing refrigerant systems it is crucial to understand precisely how these change during use; for instance knowing the temperature dependence of density will inform

pipe-sizing requirements. It is therefore of particular importance to investigate HFOs with respect to their temperature dependence of vapour pressure and vapour and liquid density.

Equations of state (EoS) have traditionally been used to simulate this behaviour and obtain thermodynamic properties of interest, and so are of relevance in many academic and industrial fields. Significant developments in the field have been seen in recent years; of note are the statistical associating fluid theory (SAFT)-derived EoS methods. Along with EoS, various correlations also see frequent use. However, EoS and correlations have some drawbacks; they can be difficult to set up and frequently require experimental data (i.e. critical point) as inputs. Many EoS also require the use of molecule-specific constants, or are only applicable to certain families of compounds (Gomez-Nieto et al., 1977). Correlations do not always provide all required outputs and underlying assumptions can limit applicability (Brown, 2012). Other issues with these methods are that they do not always predict both vapour and liquid behaviour (Sanjari et al., 2013).

An emerging area of research involves the use of machine learning techniques to correlate any number of properties of interest. Mondejar et al. (2017) recently demonstrated that machine learning can be used to develop an artificial neural network (ANN) to model certain thermophysical properties, which out performs traditional group-contribution methods to predict critical data for HFOs. This study is encouraging, but did not consider the temperature dependence of thermophysical properties.

This research explores the use of machine-learning techniques to develop an ANN capable of predicting vapour-liquid equilibria (VLE) behaviour. The procedure used by the authors to develop an ANN capable of predicting the temperature dependence of vapour pressure, and liquid and vapour densities, utilising simple molecular descriptors and other ab initio inputs, is set out. Overall, the aims of this research were to demonstrate the ability of an ANN to provide predictions with a similar accuracy to simple EoS and correlation methods, with specific reference to the HFO class of refrigerants.

2 Theory and Methodology

A step-by-step overview of the approach used is as follows; this is discussed in further detail later in this section.

1. Identify a suitable set of compounds to use to produce a predictive model
2. Collect experimental parameters from the literature
3. Use these parameters with the SAFT-γMie theoretical framework to generate a set of pseudo-experimental VLE data
4. Calculate a new set of pseudo-experimental acentric factors (ω'), to match this new data
5. Perform molecular simulations in Gaussian 16 to generate relevant ab initio data (dipole moment and polarisability)
6. Separate the collated data into three subsets; training data, test data, and validation data
7. Use the training data to train an ANN to predict VLE data as a function of temperature, utilising suitable input data
8. Evaluate the performance of the ANN on test data; based on this, refine the ANN by tweaking various hyperparameters
9. Use the validation data to verify the predictive capacity of the refined model on unseen compounds
10. Explore the use of an estimated acentric factor (ω^*) to make predictions with no requirement for experimental data, on both training and test compounds

2.1 Collating a data set for use in the ANN

2.1.1 Experimental data

A literature search was conducted for experimental data on the VLE behaviour of HFOs. Unfortunately, it was found that the available data are of irregular quality and incomplete. Semi-complete data, including critical temperature and pressure, acentric factor, and liquid density, were however found for many HFOs.

These data are sufficient inputs for the SAFT-γMie EoS and so this theory was used to generate a set of pseudo-experimental data for the development of the ANN. This approach is not ideal as there may exist systemic deviations between the SAFT data and experimental data, but was the best available option given the research needs.

2.1.2 Refrigerant selection

24 HFOs of interest were chosen, based on the availability of data and potential applicability as refrigerants.

When designing refrigerant systems it is important to consider the practical concerns (i.e. flammability, toxicity) of refrigerants in the design phase, in addition to their thermophysical behaviour. Indeed it has been theorised that mono- and di-fluorinated propene isomers (HFOs with the -1252 and -1261 base) are highly flammable, whilst other isomers with the -1225 base have been identified to be toxic (Brown, 2009). Although these properties would limit the use of such compounds as a single-component refrigerant, it is possible to mitigate these undesirable quantities

through blending with a more suitable refrigerant. The aim of this research was not to find suitable refrigerants, but to demonstrate the applicability of ANNs to predict the thermophysical properties of HFOs. These compounds were therefore included in the data set.

2.1.3 SAFT-γMie

The SAFT-γMie EoS is widely accepted as providing a reasonable approximation of real-life VLE behaviour, and has been shown to accurately model the vapour and liquid behaviour of other fluids with a low requirement for experimental data(Papaioannou et al., 2016; Rohani et al., 2011). This model was used to generate pseudo-experimental data to be used in the development of an ANN in place of experimental data.

This theory models molecules as chains of spherical beads, where each chain has an associated set of force-field parameters, which are used in a molecular simulation. In this study, the approach described by Mejía et al. (2014) was used, whereby a molecule is represented by a chain of m identical beads and force-field parameters are then calculated using the M&M correlation (developed by the same authors). The Bottled SAFT tool (Ervik et al., 2016) was used for ease, which takes values for critical temperature, pressure, and density, acentric factor and liquid density at $T = 0.7 * T_c$ to predict force-field parameters. Where this final parameter was not available, an estimation was calculated using the Rackett equation (Yamada et al., 1973).

The resulting force-field parameters were then used to run the SAFT-γMie EoS in Python, in order to generate theoretical values for critical point and VLE data as a function of temperature. Liquid and vapour volume curves were converted to density values using the molecular weight, and vapour pressure and temperature were converted into reduced terms using the critical pressure and temperature respectively, as predicted by SAFT.

Only the region of $0.5 * Tc < T < 0.9 * Tc$ was considered, as behaviour near the critical point is hard to model and rate of change is very small at lower temperatures. The temperature range used was generated according to a reverse log-scale, in order to give a higher weighting in the training process to accurately capturing the more irregular VLE behaviour seen towards higher temperatures. 100 data points were obtained for each compound, which is a sufficiently large data set to meaningfully train an ANN.

Critical point predictions were compared to experimental values. The deviation found was used to provide an indication of accuracy of the generated data and helped to decide on the number of beads to use in the molecular simulation. For almost all compounds this was the minimum value of 2. Using this optimal number of beads resulted in a mean error of 9.06% for the critical point of the HFOs considered, providing an estimate for the expected error in VLE pseudo-experimental data.

$$\omega = -log(\frac{P_{v_{T=0.7T_c}}}{P_c} - 1) \tag{1}$$

The acentric factor, defined in equation 1, simply represents a point on a vapour pressure curve, which means that the original acentric factors used as EoS inputs are no longer consistent with the new pseudo-experimental data. Using the vapour pressure curves generated by SAFT and equation 1, new acentric factors were back-calculated for each compound, denoted ω', which were used to train the ANN. These were on average 4.77% different from the experimental values

2.1.4 Electronic structure modelling

Gaussian 16 was used to optimise molecular structure and compute ab initio molecular parameters, namely the polarisibility and dipole moment, for each HFO. This computation was carried out using a density functional theory approach, with the B3LYP functional form and the 6-31G basis set. This is computationally efficient and is generally accepted as being valid in most common use cases (Tirado-Rives et al., 2008).

2.2 Building and training an ANN

An ANN is a series of connected nodes, each of which has an input x, an output y, and an associated bias, b. Each pair of connected nodes has an associated weight, w. Nodes are arranged into layers, and the model begins and ends with an input and output layer which respectively contain the same number of nodes as the inputs and outputs to the model. A network also has a number of hidden layers where tensor calculations take place to transform inputs into outputs. A non-linear activation function, $f(x)$, is applied at all hidden layer nodes, to increase ANN complexity and allow it to better handle more complex non-linear trends. The activation function is generally some kind of non-linear function, such as $tanh(x)$. The output of a node (with the exception of the input nodes) is calculated as follows:

$$x_j = \sum_i w_{ij} + b_j; \tag{2}$$

$$y_j = f(x_j) \tag{3}$$

where subscript i represents the set of preceding nodes and j the node for which the output is being calculated.

An ANN can be used to correlate complex sets of data. The inputs and outputs (sometimes known as features and labels) for the ANN used here are as follows:

Inputs:

- Reduced temperature, T/Tc
- Molecular weight, Mw

- Number of carbon atoms, n_C
- Number of fluorine atoms, n_F
- Acentric factor, ω
- Polarisability, α
- Dipole moment, m

Outputs (as a function of temperature):

- Liquid density, in the form $log(\frac{P}{P_c})$
- Liquid density, in the form $log(\rho_{liq})$
- Vapour density, in the form $log(\rho_{vap})$

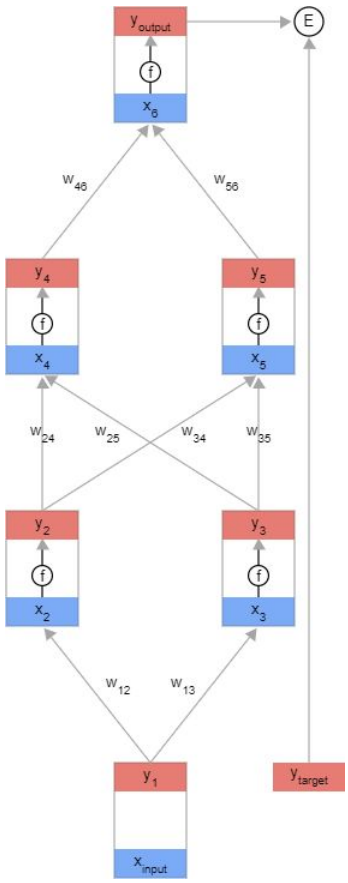


Figure 1: A representation of a simple ANN (Google, 2018)

Before the training process, which involves the optimisation of the various weights and biases of the ANN to fit the ANN's outputs to the target outputs in the training data, the data set was split into three subsets: the training set, test set, and validation set. This validation split was chosen in a non-random fashion, such that all the data for R-1132(Z), R-1234zc, and R-1336mzz(Z) were moved here. These three validation compounds were selected due to their pronounced differences in molecular structure and critical point. The remaining data were randomly split into training and test data in a 4:1 ratio, such that both these data sets included a sufficiently diverse range of data.

In addition, data was normalised by adjustment and scaling according to the training data mean and standard deviation. Throughout the study, inputs and outputs to the ANN therefore had to be normalised and denormalised as required. The outputs were provided in log format, to prevent the model from predicting non-physical negative values for pressure or density.

A back-propagation algorithm was used for the training process; this was implemented in Python using both the Keras and Tensorflow libraries in tandem, and involved the following steps:

1. Weights and biases are assigned random values between 0 and 1
2. The training data are shuffled and separated into a set number of batches
3. For a given batch, the features are passed through the model, to produce a set of label predictions
4. These predictions are compared with the corresponding target data; to do this a loss function is used as a fit metric
5. The derivative of the loss function with respect to each weight is calculated, and based on this is used to update each weight, according to the following equation:
$$w_{ij} = w_{ij} - \alpha \cdot \frac{dE}{dw_{ij}} \tag{4}$$
where α is a chosen learning rate
6. This process (from step 3 to the previous step), is repeated for each batch, until a complete pass of the training data has been made; this is known as an epoch
7. This complete pass (from step 2 to the previous step) is repeated for a set number of epochs

The mean squared error, $\sum (y_{output} - y_{target})^2$, was used as the loss function in this study. The structure of the model (number of hidden layers, number of nodes per layer, activation function, optimiser), learning rate, batch size (and thus number of batches) and number of epochs are known as the model's hyperparameters, which were manually optimised.

To evaluate the ANN's predictive capacity, two factors were considered; the accuracy of the model in fitting the training data, and the ability to generalise and predict test data accurately. This is measured using both the loss function evaluated on both sets of data, in addition to other fit metrics, including values of mean absolute relative error and R^2. It is important to have a model capable of fitting the test data no worse than the training data, so an aim

when training was to minimise test error (or equivalently, maximise test accuracy). Poor performance on test data compared with training data performance is indicative of a common problem in machine learning known as overfitting, whereby the model's ability to adapt to new data has been limited by minimising loss on training data.

To balance these factors, the model hyperparameters were optimised by the researchers through a trial and improvement process in order to minimise the value of the loss function for both the training and test data. Following this, the final model was tested on the entirely unseen validation data set to evaluate how the model could predict data for new compounds and verify the model's performance. In addition, the model performance was tested comprehensively on the entire data set, and fit parameters were tabulated.

2.3 Making predictions using an iterative acentric factor approach

Currently, the acentric factor is the only input parameter that cannot be readily deduced for any HFO; other parameters are given solely based on molecular structure, or easily generated from a Gaussian simulation. When considering new compounds, there may be no suitable experimental data to calculate this experimental point. It is therefore necessary to develop a method to approximate this acentric factor.

There have been efforts to approximate acentric factors without critical point data, however these methods have been restricted to certain families of compounds, or utilise lengthy group contribution methods (Hoshino et al., 1982). Other methods to find the acentric involve protracted or intensive processes and lack universal applicability or the ability to distinguish between isomers (Tahami et al., 2019).

Given a suitably trained ANN and a set of molecular parameters, it is proposed that through an iterative process an acentric factor for any refrigerant can be evaluated, which can then be fed back into the model to predict VLE behaviour. This process is summarised as follows:

- Assume an initial acentric factor, ω^i, between 0 and 1
- Collate ω^i with the other molecular descriptors and ab initio inputs
- Use these data with the trained ANN to predict a vapour pressure curve
- Calculate a new acentric factor, ω^{i+1}, associated with this vapour pressure curve using equation 1
- Compare the values of ω^i & ω^{i+1}
- Repeat the process using ω^{i+1} as the initial value, until $\omega^i \approx \omega^{i+1}$
- When convergence is achieved, the approximated acentric factor, ω^*, is calculated as the average of ω^i & ω^{i+1}

Acentric factors are generally between 0 and 1 for anisotropic compounds below their critical point, which narrows the search range (Mo et al., 1976). When this estimated acentric factor has been found a new set of VLE data can be generated without utilising experimental data. This represents a new method to approximate acentric factor without using any additional experimental data (once the ANN has been trained).

2.4 Comparing ANN performance against other correlations

Many correlations exist for the approximation of vapour pressure data for a compound, against which the performance of this ANN can be evaluated. The correlations set out in this subsection were selected to compare the ANN against, as they provide vapour pressure curves without critical point data. Much like the ANN, these correlations all provide reduced vapour pressure as a function of reduced temperature after making use of an acentric factor value. These correlations have been demonstrated to make accurate predictions for refrigerants (Meghdadi Isfahani et al. (2017)), but have not yet been tested on unsaturated refrigerants such as HFOs.

To allow comparison against the ANN, these correlations were run for the three validation compounds with the back-calculated acentric factor, ω′, as this was consistent with the vapour pressure data generated by SAFT.

2.4.1 Lee-Kesler Correlation

The Lee-Kesler correlation is a well-established equation that captures vapour pressure behaviour of pure fluids between their triple and critical point (Lee et al., 1975). It is defined as follows:

$$ln(P_r) = f^a(T_r) + \omega \cdot f^b(T_r) \tag{5}$$

$$f^a = 5.92714 - \frac{6.09684}{T_r} - 1.28862ln(T_r) + 0.167347T_r^6 \tag{6}$$

$$f^b = 15.2518 - \frac{15.6875}{T_r} - 13.4721ln(T_r) + 0.43577T_r^6 \tag{7}$$

where P_r is the reduced pressure, $\frac{P}{P_C}$, T_r is the reduced temperature, $\frac{T}{T_C}$, and ω' is the back-calculated acentric factor.

2.4.2 Ambrose-Walton Correlation

The Ambrose-Walton correlation is another method to model vapour pressure behaviour (Ambrose et al., 1989), and is defined as:

$$ln(P_r) = f^a(T_r) + \omega \cdot f^b(T_r) + \omega^2 \cdot f^c(T_r) \tag{8}$$

$$f^a = \frac{-5.97616\tau + 1.29874\tau^{1.5} - 0.60394\tau^{2.5} - 1.06841\tau^5}{T_r} \tag{9}$$

$$f^b = \frac{-5.03365\tau + 1.11505\tau^{1.5} - 5.41217\tau^{2.5} - 7.46628\tau^5}{T_r} \tag{10}$$

$$f^c = \frac{-0.64771\tau + 2.41539\tau^{1.5} - 4.26979\tau^{2.5} - 3.25259\tau^5}{T_r} \tag{11}$$

Where τ is defined as $1 - T_r$

2.4.3 Sanjari et al. Correlation

The Sanjari et al. correlation takes a similar form to the Lee-Kesler correlation, and was developed specifically to model refrigerants (albeit only saturated hydrocarbon refrigerants) (Sanjari et al., 2013). It is defined:

$$ln(P_r) = f^a(T_r) + \omega \cdot f^b(T_r) + \omega^2 \cdot f^c(T_r) \tag{12}$$

$$f^a = 6.834 - \frac{5.761}{T_r} + 0.907ln(T_r) - 1.169T_r^{1.9} \tag{13}$$

$$f^b = 5.320 - \frac{28.15}{T_r} - 58.04ln(T_r) + 23.57T_r^{1.9} \tag{14}$$

$$f^c = 18.20 + \frac{16.34}{T_r} + 65.7ln(T_r) - 35.97T_r^{1.9} \tag{15}$$

3 Results and Discussion

3.1 Optimal ANN hyperparameters

In training the ANN, optimal performance was obtained when using 2 hidden layers containing 12 and 6 nodes respectively, with the *tanh* activation function associated with each node. This activation function was used as it was continuous, and proved more useful than other continuous functions such as *ELU*. A learning rate of 0.005 was applied using the Adam optimiser algorithm, and the network was trained over 1000 epochs with a batch size of 6. These hyperparameters were selected after much refinement and allowed an ANN of acceptable performance to be trained over a short period of time.

The final ANN generated based on these chosen hyperparameters was able to accurately fit the training and test data. This optimisation process constituted a major part of this research in terms of time requirement; each of the hyper-parameters had to be independently changed and the ANN re-trained at each instance.

3.2 Evaluation of the ANN to predict VLE behaviour

The accuracy of the model using the training data is shown in Figure 2, demonstrating the capacity of the trained ANN to successfully predict vapour pressure, and liquid and vapour densities for HFOs.

The relative error between the model's predictions and SAFT data is less than the mean relative error of 9% found to be associated with SAFT predictions. This indicates that further optimisation of the ANN could be counterproductive, as it could lead to overfitting on the SAFT data, which is undesirable considering this model could eventually be extended to model experimental data.

The training error can be interpreted as a rough upperbound on how the model should perform with the validation data. Figure 3 shows how the model reacts to this previously unseen data. For both vapour pressure and vapour density, the model provides an accurate fit ($R^2 = 0.996$ and $R^2 = 0.988$ respectively, as shown in Figure 2).

Poor liquid density predictions were however seen ($R^2 = 0.674$ in 2). Liquid density values have large variations between different compounds and change little even within the given temperature used in this study, in contrast with vapour pressure and density. Considering this in conjunction with the small number of compounds considered may provide an explanation for this low accuracy.

3.3 Predictive ability of ANN after estimating acentric factor

A new acentric factor (ω*) was estimated for each of the training compounds, using the method described in section 2.3. A new set of predictions was then generated for each compound using the training data, with ω^* in the place of ω' as an input to the trained model.

The effectiveness of using this approach was evaluated by computing resulting fit metrics for the test data generated using this method (Figure 2), which demonstrate that this method is capable of predicting changes to vapour pressure and density without unreasonable inaccuracy. The fit was expected to be less accurate than before, given that the model was originally trained using an acentric factor that was consistent with the data set. It can be seen that $R^2 > 0.89$ for vapour pressure and vapour density, indicating that the model is still able to predict the underlying trends in these properties some extent. The liquid density is still predicted inaccurately (Figure 3).

Observing predictions for the validation compounds in Figure 3 gave similarly encouraging results. One might note that using this estimation method, predictions of vapour and liquid density on the validation data appear marginally improved compared to the training and test data set, however this is purely coincidental.

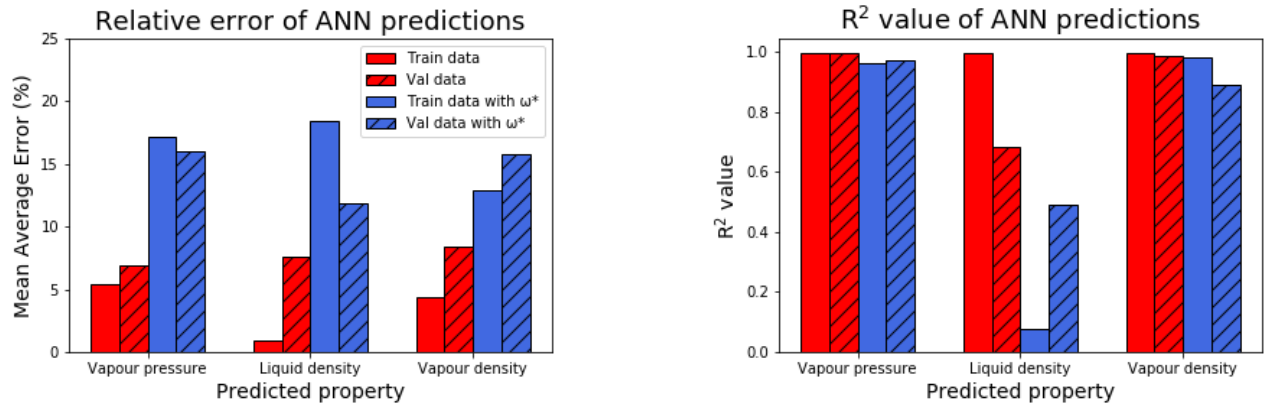


Figure 2: Evaluation of ANN predictions by two different fit metrics

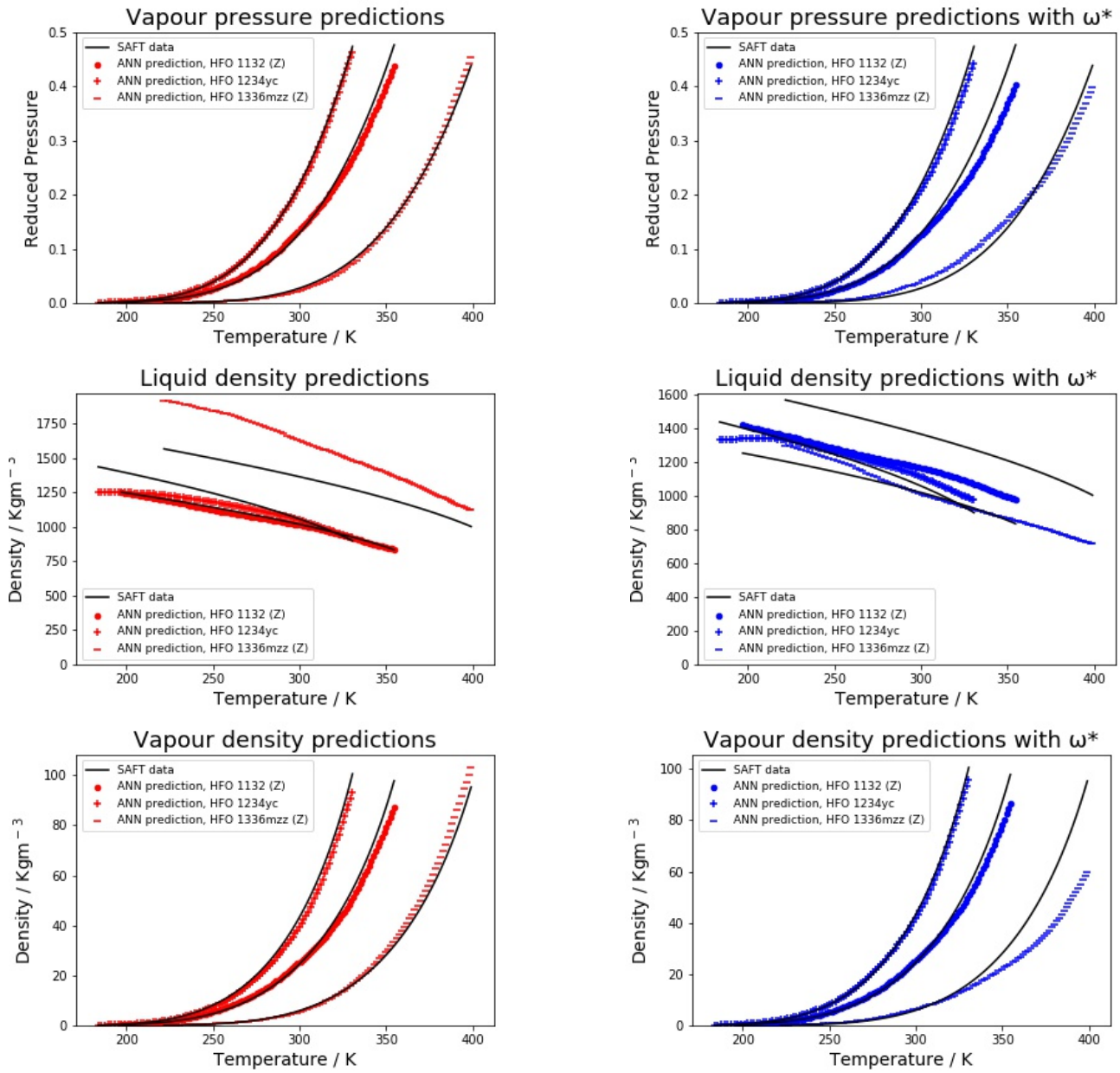


Figure 3: Comparison between SAFT data and ANN predictions using train & test data. The independent variable was shown here as temperature to allow curves to be distinguished, whereas the model worked with reduced pressure.

Table 1: Relative errors between different correlations and ANN predictions, against vapour pressure data generated from SAFT

Mean relative error (%)	Literature correlations			Developed in this study		
Refrigerant	Lee-Kesler	Ambrose-Walton	Sanjari et al.	ANN	ANN with ω*	ANN without ω
R 1132 (Z)	1.07	0.92	0.63	8.48	11.9	5.80
R 1234yc	3.25	4.09	4.81	4.04	3.97	5.49
R 1336mzz (Z)	1.51	1.67	1.59	8.24	32.8	42.6
Average	1.94	2.23	2.34	6.92	16.23	17.9

By using this estimated acentric factor method, a model has been developed that can be used without any experimental data. The model now only requires simple molecular descriptors (given when defining the HFO), ab initio quantum chemistry data, and this estimated value for acentric factor.

3.4 Performance of ANN versus other vapour pressure correlations

The performance of the previously trained ANN is evaluated in Table 1 against the Lee-Kasker, Ambrose-Walton and Sanjari et al. correlations for vapour pressure. Using the three validation compounds with the ANN showed that the literature correlations can predict pseudo-experimental data for vapour pressure more accurately than the ANN developed here. The mean relative error of the training data was 5.37% (Figure 2), indicating that existing correlations are generally more accurate than this ANN.

The ANN was slightly more accurate than the Ambrose-Walton and Sanjari et al. correlations for the vapour pressure of R-1234yc, showing the potential of the ANN. If a more complex ANN were to be trained, with a larger data set and possibly more input parameters, this ANN may be capable of predicting VLE data accurately for many different families of molecule. This would provide a clear advantage over the correlations used here, which were specifically developed to accurately model a certain class of compounds (although as shown here they are capable of predicting the vapour pressure of HFOs with reasonable accuracy).

3.5 Predictions for new HFOs

To evaluate how the iterative method developed here could estimate the acentric factor and approximate the vapour pressure curve of new HFOs, the three validation compounds were evaluated through this method. It is shown in Table 1 that the vapour pressure curve for an unknown refrigerant could be predicted slightly more accurately by using ω^*, than by using this ANN without a value for ω. Given that any small deviations in the vapour pressure curve exacerbate the errors in calculating ω^*, and that values for ω change little within families of compounds, it can be understood why using ω^* this way will not results in especially accurate predictions.The average error when using ω^* is nonetheless under 20%, demonstrating the suitability of this method as a first approximation when no other data are available.

It has been shown that there exist better correlations to generate vapour pressure curves than the ANN in its current form. For predicting the vapour pressure properties of new HFOs, using literature correlations with a suitable value for ω would likely be more accurate than using an ANN with ω^*. These correlations still require a value for acentric factor, and this is difficult due to the limited availability of experimental data for many HFOs, as well as the limitations of many approximation models as discussed in section 2.3. The values for ω^* might therefore be useful in conjunction with one of the above correlations, to provide an improved first approximation for vapour pressure curves for new HFOs.

4 Conclusions and Recommendations

This research shown here has demonstrated that ANNs are capable of predicting the temperature dependence of vapour pressure and density for HFOs. The model developed in this work displayed reasonable predictive capacity for this family of refrigerants, as shown by its ability to capture the underlying trends for 3 unseen validation compounds. The accuracy of the ANN is encouraging, however existing correlations using similar input data were shown to provide more accurate predictions.

A method to predict the VLE properties was explored, which utilised an iterative approach to estimate the acentric factor of a molecule in conjunction with a suitably trained ANN. This method has been shown to be capable of providing a first approximation to the vapour pressure and density of new HFOs for a range of temperatures (in reduced terms), without the need for any experimental data.

4.1 Research limitations

The major limitations of this research stem from the lack of available data on HFOs; a data set of only 24 compounds was used, of which 3 compounds were set aside to be used

as validation data. 20% of the remaining data was used as test data, leaving a relatively small training data set and meaning that any outlying data is able to significantly skew the model parameters.

It should be made clear that the final model developed in this research work was trained using a pseudo-experimental data set, and so the predictions produced are likely to deviate from reality. In generating SAFT data, the Rackett equation was used to estimate the required liquid density point when experimental data was unavailable. This added another layer of uncertainty and likely contributed to the ANN's failure to accurately predict liquid density. A homonuclear bead modelling approach was used in the SAFT-γMie EoS, although heteronuclear beads may have been more appropriate for often-asymmetrical HFOs.

In its current form, the ANN uses vapour pressure and temperature in reduced terms. For predictions to be meaningful, it would therefore be necessary to find critical point data, either from, experimental measurement or an estimation using an EoS or correlation.

4.2 Future research scope

Future research in this area should consider the use of a Gaussian process (Deisenroth et al., 2015), whereby an accurate model can be trained using limited data in a similar machine-learning process. The model in this research could enhanced by combining this ANN with a similar ANN able to predict critical temperature and pressure (such as the model developed by Mondejar et al. (2017)), allowing this model to provide a full replacement to a conventional EoS or correlation. The ANN can also be easily improved simply by using a wider range of experimental data.

Despite the limitations described in the previous subsection, in combination with existing research in the field (Mondejar et al., 2017), this paper presents a strong case for further investigation of the use of ANNs in fitting and predicting VLE behaviour. With a sufficiently large data set of accurate experimental data, many of the problems encountered in this study can be avoided.

5 References

Ambrose, D et al. (1989). "Vapour pressures up to their critical temperatures of normal alkanes and I-alkanols". en. In: p. 9.

Amendment to the Montreal Protocol on Substances that Deplete the Ozone Layer (Sept. 2016). UN Treaty. United Nations. URL: https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-2-f&chapter=27&clang=_en.

Avendaño, Carlos et al. (Feb. 2013). "SAFT-γForce Field for the Simulation of Molecular Fluids: 2. Coarse-Grained Models of Greenhouse Gases, Refrigerants, and Long Alkanes". In: *The Journal of Physical Chemistry B* 117.9, pp. 2717–2733. DOI: 10.1021/jp306442b.

Brown J. Steven; Polonara, Fabio; Di Nicola Giovanni; Fedele Laura; Bobbo Sergio (2012). "Vapor Pressure of Hydrofluoroolefins: Critical Review of Experimental Data and Models". In: *International Refrigeration and Air Conditioning Conference*. 1316. URL: https://docs.lib.purdue.edu/iracc/1316/.

Brown, J. Steven (Aug. 2009). "HFOs: New, Low Global Warming Potential Refrigerants". English. In: *ASHRAE Journal; New York* 51.8, pp. 22–24,26,28–29. ISSN: 00012491. URL: http://search.proquest.com/docview/220448891/abstract/310AEA9051064501PQ/1 (visited on 10/23/2019).

Brown, J. Steven et al. (Nov. 2016). "Low-GWP refrigerants". In: *Science and Technology for the Built Environment* 22.8, pp. 1075–1076. ISSN: 2374-4731. DOI: 10.1080/23744731.2016.1230981.

Deisenroth, Marc Peter et al. (2015). "Distributed Gaussian Processes". In: *Proceedings of the 32nd International Conference on Machine Learning*. URL: http://proceedings.mlr.press/v37/deisenroth15.pdf/.

Ervik, Åsmund et al. (Aug. 2016). "Bottled SAFT: A Web App Providing SAFT-γMie Force Field Parameters for Thousands of Molecular Fluids". In: *Journal of Chemical Information and Modeling* 56.9, pp. 1609–1614. DOI: 10.1021/acs.jcim.6b00149.

Frisch, M. J. et al. (2016). *Gaussian 16, Gaussian, Inc., Wallingford CT*.

Gomez-Nieto, Mateo et al. (May 1977). "A New Vapor Pressure Equation and Its Application to Normal Alkanes". en. In: *Industrial & Engineering Chemistry Fundamentals* 16.2, pp. 254–259. ISSN: 0196-4313, 1541-4833. DOI: 10.1021/i160062a014.

Google, ed. (2018). *Machine Learning Crash Course with TensorFlow APIs*. URL: https://developers.google.com/machine-learning/crash-course/.

Hoshino, D. et al. (1982). "Prediction of Acentric Factor of Alkanes by the Group Contribution Method". In: *Journal of Chemical Engineering of Japan* 15.2, pp. 153–155. DOI: 10.1252/jcej.15.153.

Kyoto Protocol to the United Nations Framework Convention on Climate Change (Dec. 1997). UN Treaty. United Nations. URL: https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-7-a&chapter=27&clang=_en.

Lee, Byung Ik et al. (1975). "A generalized thermodynamic correlation based on three-parameter corresponding states". In: *AIChE Journal* 21.3, pp. 510–527.

Meghdadi Isfahani, Amir Homayoon et al. (2017). "Empirical Correlations and an Artificial Neural Network

Approach to Estimate Saturated Vapor Pressure of Refrigerants". In: *Physical Chemistry Research* 5.2, pp. 281–292. ISSN: 2322-5521. DOI: 10.22036/pcr.2017.41111.

Mejía, Andrés et al. (2014). "Force fields for coarse-grained molecular simulations from a corresponding states correlation". In: *Industrial & Engineering Chemistry Research* 53.10, pp. 4131–4141.

Mo, Kingtse C. et al. (1976). "Perturbation theory for anisotropic fluid transport coefficients". In: *The Journal of Chemical Physics* 65.9, pp. 3715–3719. DOI: 10.1063/1.433560.

Mondejar, M.E. et al. (2017). "Prediction of properties of new halogenated olefins using two group contribution approaches". In: *Fluid Phase Equilibria* 433, pp. 79–96. DOI: 10.1016/j.fluid.2016.10.020.

Morgado, Pedro et al. (Sept. 2016). "SAFT-γforce field for the simulation of molecular fluids: 8. Hetero-segmented coarse-grained models of perfluoroalkylalkanes assessed with new vapour–liquid interfacial tension data". In: *Molecular Physics* 114.18, pp. 2597–2614. DOI: 10.1080/00268976.2016.1218077.

Papaioannou, Vasileios et al. (2014). "Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments". In: *The Journal of Chemical Physics* 140.5, pp. –.

Papaioannou, Vasileios et al. (May 2016). "Application of the SAFT-γMie group contribution equation of state to fluids of relevance to the oil and gas industry". In: *Fluid Phase Equilibria* 416, pp. 104–119. DOI: 10.1016/j.fluid.2015.12.041.

Rohani, Ali Asghar et al. (Mar. 2011). "Comparison between the artificial neural network system and SAFT equation in obtaining vapor pressure and liquid density of pure alcohols". In: *Expert Systems with Applications* 38.3, pp. 1738–1747. DOI: 10.1016/j.eswa.2010.07.099.

Sanjari, Ehsan et al. (June 2013). "An accurate generalized model for predict vapor pressure of refrigerants". en. In: *International Journal of Refrigeration* 36.4, pp. 1327–1332. ISSN: 01407007. DOI: 10.1016/j.ijrefrig.2013.01.007.

Secretariat, UNEP Ozone (2003). *Handbook for the International Treaties for the Protection of the Ozone Layer*.

Tahami, Saba et al. (Nov. 2019). "Estimation of the acentric factor of organic compounds via a new group contribution method". In: *Fluid Phase Equilibria* 499, p. 112246. ISSN: 0378-3812. DOI: 10.1016/j.fluid.2019.112246.

Tirado-Rives, Julian et al. (Feb. 2008). "Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules". In: *Journal of Chemical Theory and Computation* 4.2, pp. 297–306. DOI: 10.1021/ct700248k.

Yamada, Tomoyoshi et al. (Apr. 1973). "saturated liquid molar volumes. Rackett equation". In: *Journal of Chemical & Engineering Data* 18.2, pp. 234–236. ISSN: 0021-9568. DOI: 10.1021/je60057a006.

6 Appendices

6.1 Refrigerants considered

Table 2: Refrigerants included in the training and test data set

Refrigerant	Systematic Name
R-1114	tetrafluoroethene
R 1123	trifluoroethene
R 1132a	1,1-difluoroethene
R 1132 (E)	trans-1,2-difluoroethene
R 1132 (Z)	cis-1,2-difluoroethene
R 1141	fluoroethene
R 1216yc	1,1,2,3,3,3-hexafluoropropene
R 1225yc	1,1,2,3,3-Pentafluoropropene
R 1225ye (E)	trans-1,2,3,3,3-Pentafluoropropene
R 1225ye (Z)	cis-1,2,3,3,3-Pentafluoropropene
R 1225zc	1,1,3,3,3-Pentafluoropropene
R 1234yc	1,1,2,3-Tetrafluoropropene
R 1234ye (E)	trans-1,2,3,3-Tetrafluoropropene
R 1234ye (Z)	cis-1,2,3,3-Tetrafluoropropene
R 1234yf	2,3,3,3-Tetrafluoropropene
R 1234zc	1,1,3,3-Tetrafluoropropene
R 1234ze (E)	trans-1,3,3,3-Tetrafluoropropene
R 1234ze (Z)	cis-1,3,3,3-Tetrafluoropropene
R 1243zf	3,3,3-trifluoropropene
R 1261yf	2-fluoro-1-propene
R 1261ze (E)	trans-1-fluoro-1-propene
R 1261ze (Z)	cis-1-fluoro-1-propene
R 1261zf	3-fluoro-1-propene
R 1336mzz (E)	trans-1,1,1,4,4,4-hexafluoro-2-butene

Factors Affecting Protein Trans-splicing using Split Inteins

Derek Lam and Alexa Vasey
Department of Chemical Engineering, Imperial College London, U.K.

Abstract

Bispecific monoclonal antibodies (BsMAbs) have the potential to effectively treat cancer but they are notoriously difficult to produce. One method to produce BsMAbs is through the use of inteins which overcomes the light chain pairing problem encountered in traditional BsMAbs production. Inteins have the ability to covalently ligate proteins together when coexpressed in trans, giving a modular approach in building BsMAbs. This paper focuses on understanding previously neglected factors that influence protein trans-splicing and raises questions about selecting appropriate model proteins, something that has not been explored enough in other literature. The findings show that dithiothreitol (DTT) and EDTA concentrations could either help or hinder the reactions, depending on the intein used. This suggests that one set of trans-splicing conditions may not be optimal for all inteins. It was also shown that EDTA significantly improves the half life of DTT, slowing down the conversion to its oxidised form. This maintains the reducing capabilities of DTT over longer periods, without the need for increasing the initial DTT concentration which could have adverse effects in many processes, such as antibody production.

Keywords: DTT, EDTA, model protein, protein trans-splicing, reduction, split-inteins

Introduction

Bispecifc monoclonal antibodies (BsMAbs) are becoming a promising option to treat many diseases such as cancer. They have two different antigen binding sites which can, for example, bind onto tumour cells and T-cells simultaneously.[1] BsMAbs are hard to synthesise using traditional methods due to the chain association issue. A modular approach can be taken to build BsMAbs through the use of split inteins to overcome this issue.[2]

Split inteins are proteins that have two separate proteins attached to either end (exteins). They have the ability to excise themselves and covalenty bond both exteins together.[3,4,5] This has many different applications such as protein purification, protein modification and intein based biosensors.[6,7] Ramirez et al. demonstrated the ability to engineer split-inteins from E. coli expression.[8] Before trans-splicing occurs, the inteins are primed using a reducing agent such as dithiothreitol (DTT). This reduces any disulphide bonds formed due to the presence of cysteines in the intein.[9] The stability of DTT is much greater in crystalline form than in solution.[10] This means that DTT can significantly oxidise over a 24 hour reaction, decreasing it's reducing capacity, causing the inteins to revert back to their oxidised state. This could affect both reaction rates and yields. In some circumstances, increasing the initial DTT concentration is not a viable option as it could affect key disulphide bonds present in the exteins, such as in the context of antibodies. Lower concentrations of reducing agent are therefore preferred.

This paper explores whether DTT has any limitations on the trans-splicing reaction, ways to decrease the decomposition of DTT and the effect of these methods on individual trans-splicing reaction yields. It also seeks to raise awareness about testing the suitability of model proteins used in research.

Background

Carvajal-Vallejos et al. showed that specific split inteins (GP41-1, GP41-8, NRDJ-1 and IMPDH) showed extremely high reaction rates and efficiencies. GP41-1 showed reaction rates that were 10 times higher than *Npu* DnaE.[11] However when the split inteins used in Carvajal-Vallejos et al. were tested using green and red fluorescent proteins (GFP and RFP), the reaction rates and yields were significantly lower than the ones reported. Papers such as Wright et al. have also reported not reaching these exceptionally high yields either.[12] This is an important point since other papers such as Bowen et al. build their hypotheses assuming the rates and yields in Carvajal-Vallejos et al. are accurate.[13] If this were not the case, then many ideas would need to be reformulated, hence this paper set out to optimise the splicing conditions of the inteins used in this study.

One of the differences between this study and Carvajal-Vallejos et al. was the supplementation of EDTA to the final buffer. The effect of which on the split inteins and the protein trans-splicing reaction has not been tested. Another difference was the model protein attached to the intein. Cheryian et al. and Zettler et al. show that the trans-splicing can be dependant on the extein.[14,15] In Carvajal-Vallejos et al., Thioredoxin (TRX) was used as the C extein, while in this paper, RFP was used. A study using an insulin reduction assay shows that TRX supplemented DTT solutions accelerate DTT reduction of disulphide bonds by around 20 fold.[16]

Most inteins require a reducing environment to facilitate protein trans-splicing.[9,17] Therefore it begs the question of what effects TRX would have on the rates/yields of the trans-splicing reaction in the context of its enhanced reducing capacity. If proven true, it would therefore be inaccurate to report rate constants for split inteins in the presence of TRX when it is unlikely that split inteins would be used in applications where TRX is the C extein. This naturally gives rise to the question of whether the same yield as the Carvajal-Vallejos et al. study could be achieved from DTT alone.

Materials and methods

Plasmids - The pCK304 expression vector encoding the chosen N and C split inteins (NRDJ-1, IMPDH, GP41-1 and GP41-8) fused to sfGFP and mScarlett was provided by the Polizzi lab.

Expression - For each intein, 5μl of ampicillin at 100mg/ml, 1000x the working concentration, was added to 5ml of LB broth in a 15ml falcon tube. E. coli cultures held at -20°C were added to the LB broth/ampicillin mixture. The cultures were left in a 37°C shaking incubator overnight. 5ml of culture was added to 100ml of LB broth supplemented with ampicillin. This was then put in incubation at 30°C for 2 hours. Cultures were induced with L-rhamnose to a final concentration of 1mg/ml to switch the E. coli from it's off state and start producing GFP/RFP with inteins attached. The cultures were placed back into incubators at 30°C for 12 hours.

His-affinity purification - The incubated cultures were centrifuged at 6000g for 5 minutes causing the E. coli to settle to the bottom of the tube. The supernatant was then discarded. The cells were resuspended in 10ml of a buffer containing 50mM Na_2HPO_4, 25mM imidazole and 300mM NaCl at pH8. The suspension was placed in ice while preparing for sonication. The cells were lysed via sonication in an ice-water mixture to dissipate heat. The soluble fraction was obtained after centrifugation at 13000g for 50 minutes. The supernatant containing the proteins was applied to 0.25ml of Ni-NTA agarose beads were added to capture the protein. This was left in an ice bath and placed on a rocker shaker for 1.5 hours. The

agarose beads were filtered out by passing it through a 5ml column. 4ml of wash buffer containing 50mM Na_2HPO_4, 50mM imidazole and 300mM NaCl at pH8 was added to wash the beads. Proteins were eluted using an elution buffer (250mM imidazole, 300mM NaCl, 50mM Na_2HPO_4 at pH8). Three elution fractions of 0.5ml were collected in separate Eppendorf tubes. The proteins were then dialysed with a 50mM TrisHCl 0.3mM NaCl 10% (v/v) glycerol splicing buffer at pH 7 at 5°C in a stirred beaker overnight.The proteins were subsequently filtered through a 200nm Sartorius Minisart filter.

Determining concentration - Concentrations were determined using a Bradford assay interpolating from a bovine serum albumin (BSA) standard . 5μl of each standard/sample was coincubated with 250μl of Bradford reagent at room temperature for 10 minutes in wells on a Costar96 plate. In the presence of proteins, a colorimetric shift was observed. The lid is put on the plate, making sure there are no bubbles in the wells first as this will affect the readings. The position of the standards, blanks (pure reagent) and proteins were added to CLARIOstar. Absorbance values at a wavelength of 595nm were then measured.

Trans-splicing - The proteins were diluted to 10μM reaction concentrations and DTT was added to a final concentration of 2mM. The proteins were then coincubated to start the reaction. The reaction was monitored over 24 hours with samples taken from the reacting mixture and quenched with 5x SDS loading dye (5x solution of 250mM TrisHCL, 10% SDS, 30%(v/v) glycerol, 10mM DTT, 0.05%(w/v) Bromophenol blue, pH 6.8) and boiled at 95°C for 2 minutes using an Applied Biosystems Veriti 96-Well Thermal Cycler.

Cross Reactivity - The inteins were coincubated in orthogonal and non-orthogonal pairs in a 4 x 4 matrix at 5μM final concentration. They were left to react for 24 hours before being quenched with SDS-PAGE sample buffer and boiled in preparation to be loaded onto 12.5% gels.

Yield analysis - 12.5% acrylamide gels were made to run the samples from the GFP/RFP reactions. 8μl of PageRuler Prestained Protein Ladder 10 to 180kDa was loaded in the first lane to act as a reference for sizes. 4μl of each sample was loaded into the next wells of the gel. The gels were run at 140V for 120 minutes. Gels were then removed and washed with distilled water 3 times for 5 minutes to remove SDS. After the washing step, gels were stained with Simply Blue Safestain for 1 hour, before being destained with distilled water, imaged and analysed with ImageJ software.[18] Relative band intensities (BI) were measured through the ImageJ program. Yield was calculated from the depletion of GFP + intein between the first and last time point, measured by band intensities as in equation 1.

$$Yield = \frac{BI_{t=0} - BI_{t=24}}{BI_{t=24}} \tag{1}$$

Oxidised DTT standard - The oxidised DTT standard curve was produced using a Bradford assay. Specific DTT concentrations were coincubated with Bradford reagent for 10 minutes. This was then loaded into the CLARIOstar to retrieve the absorbance results.

DTT minimum working concentrations - The minimum working concentrations were calculated by using an array of decreasing DTT concentrations for the trans-splicing reaction. Serial dilutions were made of DTT from 2mM to 0.016mM. The trans-splicing reactions were run with 2μl of each DTT dilution for each intein. They were then run on 12.5% acrylamide gels following the same methodology in kinetic analysis. However, densitometric analysis was conducted on the spliced product instead of reactants, to determine band intensities. Relative yield was found by assuming the 2mM DTT

concentration was the maximum yield. The DTT minimum working concentration can be found by plotting the relative yield against the concentration of DTT.

Half life of DTT - The half life of DTT was found by preparing a 2mM solution in the splicing buffer and leaving it for 12 hours. This was then loaded into the CLARIOstar along with a blank to take the absorbance readings. From the standard curve of oxidised DTT the blank adjusted concentration could be found. Subsequently, the change in oxidised DTT over the 12 hours was found, giving the change in DTT concentration. This was then used to calculate the half life of DTT.

Results and discussion

Cross reactivity - Intein orthogonality has been seen to vary, so cross reactivity was tested.[19] Table 1 shows the results of a cross-reactivity assay which implies that there is no detectable cross-reactivity. Therefore, the inteins are not active in their non-native configurations. This is an important consideration for multi-part assemblies, such as producing BsMAbs, to ensure the correct configuration is achieved. This is a result that has previously been suggested in literature but data has not been shown.[11]

Table 1: A matrix showing whether a spliced product formed after cross reacting intein termini at 5μM final concentration for 24 hours

	1C	2C	3C	4C
1N	Yes	-	-	-
2N	-	Yes	-	-
3N	-	-	Yes	-
4N	-	-	-	Yes

Where: 1 = NRDJ-1, 2 = IMPDH, 3 = GP41-1, 4 = GP41-8.

Trans-splicing reaction - Figure 1 shows an example of the SDS-PAGE gels obtained from the protein splicing reactions. A spliced product (SP) can be seen forming above the 55KDa band, while the GFP + intein band, just below the 40KDa band, is fading showing that the protein is being depleted. Table 2 shows that the observed yields from this experiment are significantly lower than the yields reported in Carvajal-Vallejos et al.[11] The reaction was found to have reached a plateau because the intensity of the reactant band between 18 and 24 hours did not change. This means there was no more reactant depletion over this period. The difference in yields prompted an investigation as to why this was occurring. Two explanations were suggested, the presence of TRX and DTT oxidation.

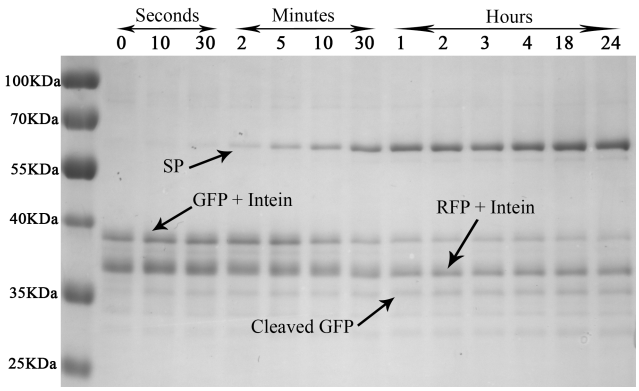


Figure 1: SDS-PAGE gel of splicing reaction using intein GP41-8 at 5μM final concentration over 24 hours

Table 2: Yields of different inteins at 5μM final concentration after 24 hours without EDTA compared to yields from Carvajal-Vallejos et al.

Intein	Experimental Yield (%)	Carvajal-Vallejos Yield (%)
IMPDH	43 ± 8	90-95
GP41-1	46 ± 4	80-90
GP41-8	67 ± 3	85-95

Rate graphs were not plotted due to the small variance in the intensity of the bands on the SDS PAGE gel at the start of the reaction. This meant that the depletion of the reactant was too small a change to detect, meaning

rates would have been unreliable given the imaging method used. NRDJ-1 also had poor band intensity in subsequent reactions after testing for cross reactivity. Therefore results related to NRDJ-1 from this point onwards have been omitted.

Oxidised DTT standard - In order to conduct analysis on study the oxidation of DTT in the splicing buffer, a calibration curve was created for a range of DTT concentrations within the experimental working range. Figure 2 shows that the relationship of absorbance and oxidised DTT was linear. Although the peak of oxidised DTT absorbance is 280nm, the measurement was taken at 310nm so that there was less interference for further analysis with the proteins.[20]

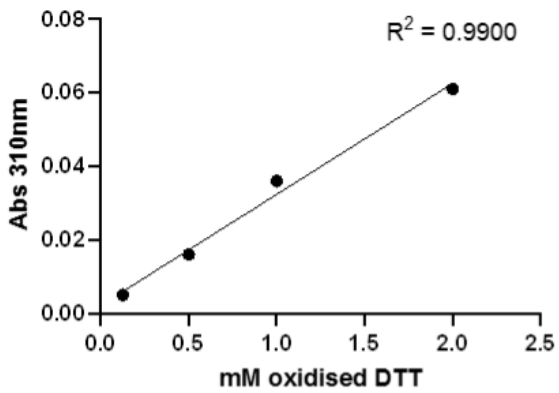


Figure 2: Standard curve for oxidation vs. absorbance at 310nm of DTT

DTT half life - The half life of DTT in the splicing buffer was found to be 5.2 hours. This means that over the course of the reaction the 2mM initial concentration decreased to just 0.08mM by 24 hours. This degradation rate could then have implications on the trans-splicing reaction.

DTT minimum working concentration - Figure 3 shows how the formation of the spliced product varies with DTT concentration for GP41-1. As the DTT concentration decreases, the spliced product band intensity decreases, disappearing at 0.12mM DTT. Figure 4 shows GP41-8 reacting under the same DTT concentrations. GP41-8 catalayses the trans-splicing reaction and produces a spliced product at all tested DTT concentrations. The spliced product band formed at 0mM DTT is slightly lower than the bands formed at higher DTT concentration. This could be an oxidised form of the spliced product, as there is no DTT to create a reducing environment. Oxidised/reduced proteins have different migrations on a SDS gel which could explain the observations.

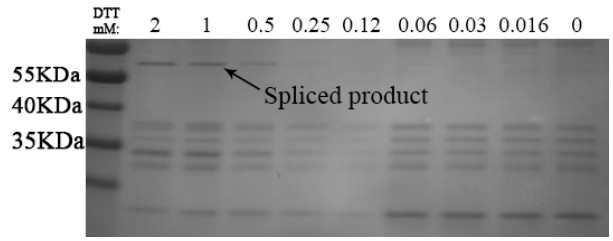


Figure 3: DTT Assay for GP41-1

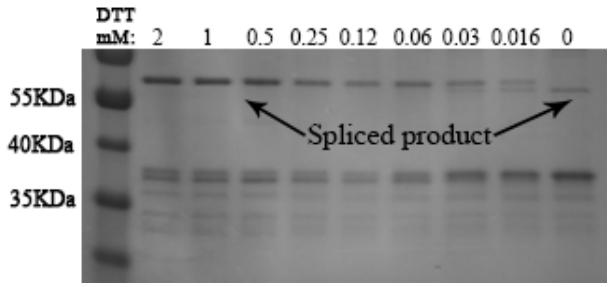


Figure 4: DTT Assay for GP41-8

Figure 5 shows IMPDH and GP41-1 having a minimum working concentration of 0.02mM and 0.17mM respectively. GP41-8 can catalyse the protein trans-splicing reaction without DTT. Collectively, these results show that split intein susceptibility to DTT differs.

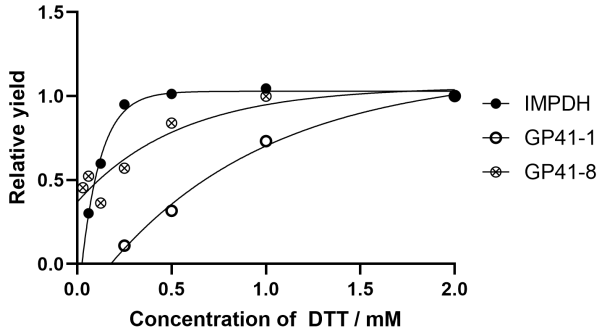


Figure 5: Relative yield of each intein vs. DTT concentration without EDTA after 24 hours

EDTA - Figure 6 shows the change in DTT oxidation over a 12 hour period in the presence of EDTA. The protective effect of EDTA on DTT concentration is consistent with findings from Stevens et al.[21] The half life of DTT when supplemented with EDTA can be increased to 117.7 hours from 5.2 hours. As an example, after 24 hours the initial 2mM DTT solution would only decrease to 1.73mM.

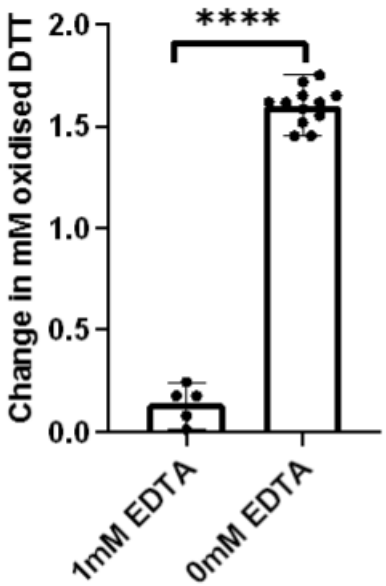


Figure 6: The effect of 1mM EDTA on 2mM DTT in splicing buffer over 12 hours. **** represents high significance (P value ≤ 0.0001)

Effect of EDTA supplementation on the minimum working DTT concentration - The DTT minimum working concentration was then repeated with the supplementation of EDTA and the relative yield vs. DTT concentration graph was reproduced. Figure 7 shows IMPDH and GP41-1 shift their minimum DTT thresholds left. This shows trans-splicing occured at lower DTT concentrations. Figure 8 shows a direct comparison for GP41-1 with and without the addition of EDTA. EDTA seems to increase the susceptibility of the proteins to DTT allowing a higher amount of it's maximum yield to be achieved at a certain DTT concentration. Nicastri et al. also observed this behaviour for *Npu* DnaE.[9] However as seen in figure 9, IMPDH has had it's susceptibility decrease slightly. The reason for this is unknown but it shows that adding EDTA does not seem to be broadly applicable to all inteins.

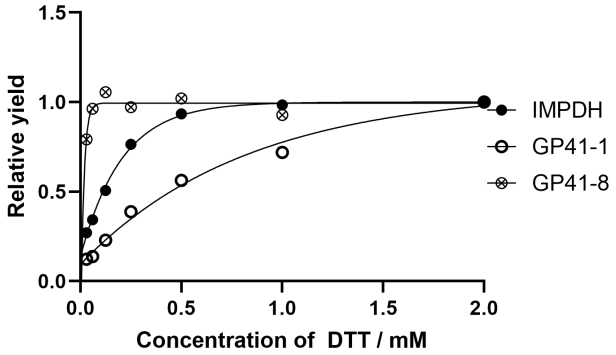


Figure 7: Relative yield of the inteins vs. DTT concentration with 1mM EDTA after 24 hours

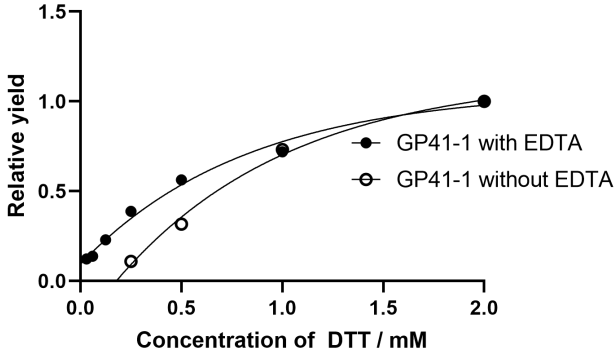


Figure 8: GP41-1's dependency on DTT with and without 1mM EDTA after 24 hours

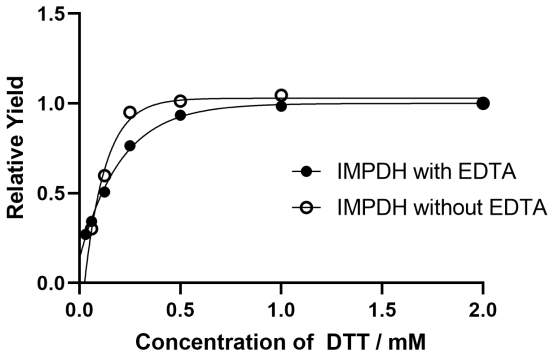


Figure 9: IMPDH's dependency on DTT with and without 1mM EDTA after 24 hours

Trans-splicing reaction with EDTA - The protein trans-splicing reaction of inteins was monitored with the addition of EDTA to look for improvement in yield. Table 3 shows the effect of the addition of EDTA on the yield at 24 hours for three split intein pairs. GP41-1 did not have a statistically significant change in yield so the effects of EDTA can not be concluded. The average yield of IMPDH slightly increased. This could be due to the increased stability of DTT over the 24 hour period, or it could be due to a beneficial effect from the presence of EDTA. In contrast, GP41-8 had a decrease in the yield. This could imply that EDTA actually has some negative effects on the trans-splicing reaction for some inteins. This could suggest that EDTA has both positive and negative effects, in which the cumulative effect varies between inteins.

Table 3: Change in each intein's yield after 24 hours with the addition of 1mM EDTA

Intein	Without EDTA Yield (%)	With EDTA Yield (%)
IMPDH	43 ± 8	57 ± 3
GP41-1*	46 ± 4	53 ± 2
GP41-8	67 ± 3	40 ± 2

*This result was not statistically significant, so no implications could be drawn.

Metal Ions - Metal ions have been shown to increase the oxidation of DTT in solution.[22] EDTA is a known chelating agent, due to the significant increase in DTT half-life after the addition of EDTA, it was suspected that metal ions may be present. After looking at the certificate of analysis on the buffer reactants it was seen that trace amounts of lead and magnesium were present in the sodium chloride.[23] Although the proteins were filtered before trans-splicing occurred, this may not have prevented these metals from entering the reaction stock which could facilitate the oxidation of DTT.

Conclusions and outlook

EDTA's effect on DTT - EDTA can be seen to increase the half life of DTT in the protein splicing buffer. This result is consistent with other literature.[21,22]

EDTA's effect on trans-splicing - EDTA appears to have a complex effect on the trans-splicing reaction which is split intein specific. For certain inteins, such as IMPDH, it might increase the yield. However for inteins such as GP41-8 a decrease in yield was seen. As such, the use of EDTA should be considered with respect to the choice of inteins. The increase in yield for IMPDH and was still not enough to reach that reported in Carvajal-Vallejos et al. and so another hypothesis is needed as to understand why.[11]

Testing with TRX - The yields reported by Carvajal-Vallejos et al. were still not reached even with the improved DTT half life. As mentioned before, TRX has a significant influence on the rate of reduction.[16] The effects of TRX on the trans-splicing reaction should therefore be investigated as this could explain the differences between the rates and yields that others, such as Wright et al., achieve and Carvajal-Vallejos et al.[11,12]

Improving analysis methods - In order to verify the effects of adding EDTA to the trans-splicing reaction, more repeats would be needed. Using a different analysis method such as HPLC, capillary electrophoresis or a different densitometric imaging software such as Quantity One (Bio-Rad) program would greatly increase the accuracy of measurements, decreasing the error in calculations.

Testing a cysteine-less split intein - A new cysteine-less split intein has been found which could potentially bypass the problem laid out in this study.[24] DTT is used to reduce disulphide bonds present in the intein which cannot form without cysteines. Therefore a cysteine-less intein would not need a reducing agent avoiding the need for DTT and EDTA.

The yield of such an intein should be tested to see if it were efficient at trans-splicing, especially in an antibody context.

Testing new reducing agents - Investigation into different reducing agents should be carried out. Other studies use TCEP as their reducing agent.[9,25] In a direct comparison by AGScientific, TCEP HCl is said to be more stable and more powerful than DTT,[26] and so could provide higher reducing powers over a longer reaction time.

Antibody fragments - As these inteins have only been tested with model proteins, they should also be tested with antibody pieces to determine their efficiency. For example, CH_2 and CH_3 could be tested.

Acknowledgements

We would like to express our thanks to Oskar Lange for his invaluable support and guidance throughout the project, as well as to all other members of the Karen Polizzi lab for their support.

References

[1] Baeuerle PA, Reinhardt C. Bispecific T-cell engaging antibodies for cancer therapy. *Cancer Research*. 2009;69(12):4941–4944. Available from: https:doi.org/10.1158/0008-5472.CAN-09-0547.

[2] Han L, Chen J, Ding K, Zong H, Xie Y, Jiang H, et al. Efficient generation of bispecific IgG antibodies by split intein mediated protein trans-splicing system OPEN. *Sci Rep*, page=8360, volume=7;Available from: https://doi.org/10.1038/s41598-017-08641-3.

[3] Saleh L, Perler FB. Protein splicing In Cis and In Trans. *The Chemical Record*. 2006;6(4):183–193. Available from: http://doi.wiley.com/10.1002/tcr.20082.

[4] Perler FB, Davis EO, Dean GE, Gimble FS, Jack WE, Neff N, et al. Protein splicing elements: Inteins and exteins - a definition of terms and recommended nomenclature. *Nucleic Acids Research*. 1994;22(7):1125–1127. Available from: https://doi.org/10.1093/nar/22.7.1125.

[5] Mills KV, Lew BM, Jiang SQ, Paulus H. Protein splicing in trans by purified N- and C-terminal fragments of the Mycobacterium tuberculosis RecA intein. *PNAS*. 1998;95(7):3543–3548. Available from: https://doi.org/10.1073/pnas.95.7.3543.

[6] Wood DW, Camarero JA. Intein applications: From protein purification and labeling to metabolic control methods. *Journal of Biological Chemistry*. 2014;289(21):14512–14519. Available from: https://doi.org/10.1074/jbc.R114.552653.

[7] Topilina NI, Mills KV. Recent advances in *in vivo* applications of intein-mediated protein splicing. *Mobile DNA*. 2014;5(1). Available from: https://doi.org/10.1186/1759-8753-5-5.

[8] Ramirez M, Valdes N, Guan D, Chen Z. Engineering split intein DnaE from Nostoc punctiforme for rapid protein purification. *Peds*. 2013;26(3). Available from: https://doi.org/10.1093/protein/gzs097.

[9] Nicastri MC, Xega K, Li L, Xie J, Wang C, Linhardt RJ, et al. Internal disulfide bond acts as a switch for intein activity. *ACS Publications*. 2013;52(34):5920–5927. Available from: https://doi.org/10.1021/bi400736c.

[10] AGScientific. *DTT (Dithiothreitol): Frequently Asked Questions*;. Available from: https://bit.ly/34fo9aS.

[11] Carvajal-Vallejos P, Pallissé R, Mootz HD, Schmidt SR. Unprecedented rates and efficiencies revealed for new natural split inteins from metagenomic sources. *The Journal of biological chemistry*. 2012;287(34):28686–96. Available from: http://doi.org/10.1074/jbc.M112.372680.

[12] Wright JN, Wong WL, Harvey JA, Garnett JA, Itzhaki LS, Main Correspondence ERG. Scalable Geometrically Designed Protein Cages Assembled via Genetically Encoded Split Inteins. *Structure*. 2019;27:776–784.e4. Available from: https://doi.org/10.1016/j.str.2019.02.005.

[13] Bowen CH, Reed TJ, Sargent CJ, Mpamo B, Galazka JM, Zhang F. Seeded Chain-Growth Polymerization of Proteins in Living Bacterial Cells. *ACS Synthetic Biology*. 2019;8:acssynbio.9b00362. Available from: https://doi.org/10.1021/acssynbio.9b00362.

[14] Cheriyan M, Pedamallu CS, Tori K, Perler F. Faster protein splicing with the nostoc punctiforme DnaE intein using non-native extein residues. *Journal of Biological Chemistry*. 2013;288(9):6202–6211. Available from: https://doi.org/10.1074/jbc.M112.433094.

[15] Zettler J, Schütz V, Mootz HD. The naturally split *Npu* DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Letters*. 2009;583(5):909–914. Available from: http://doi.org/10.1016/j.febslet.2009.02.003.

[16] Holmgren A. Thioredoxin Catalyzes the Reduction of Insulin Disulfides; 1979. Available from: http://www.jbc.org/content/254/19/9627.short.

[17] Calce E, Rosa ·, Vitale M, Scaloni A, Amodeo P, De Luca S. Air oxidation method employed for the disulfide bond formation of natural and synthetic peptides. *Amino acids*. 1983;Available from: https://doi.org/10.1007/s00726-015-1983-4.

[18] Rasband WS. *Image Processing and Analysis in Java*;. Available from: https://imagej.nih.gov/ij/.

[19] Aranko AS, Oeemig JS, Zhou D, Kajander T, Wlodawer A, Iwaï H. Structure-based engineering and comparison of novel split inteins for protein ligation. *Mol BioSyst*. 2014;10:1023–1034. Available from: http://doi.org/10.1039/C4MB00021H.

[20] Hegyi G, Kardos J, Kovács M, Málnási-Csizmadia A, Nyitray L, Pál G, et al. Introduction to Practical Biochemistry. Eötvös Loránd University; 2013. Available from: https://bit.ly/2YEbmh7.

[21] Stevens R, Stevens L, Price N. The stabilities of various thiol compounds used in protein purifications. *Biochemical Education*. 1983;11(2):70. Available from: https://doi.org/10.1016/0307-4412(83)90048-1.

[22] Getz EB, Xiao M, Chakrabarty T, Cooke R, Selvin PR. A Comparison between the Sulfhydryl Reductants Tris(2-carboxyethyl)phosphine and Dithiothreitol for Use in Protein Biochemistry. *Analytical Biochemistry*. 1999;273. Available from: https://doi.org/10.1006/abio.1999.4203.

[23] VWR. Characteristics Specifications;. Available from: https://qr.vwr.com/assetsvc/asset/en_GB/id/14520801/contents/14520801.pdf.

[24] Bhagawati M, Terhorst TME, Füsser F, Hoffmann S, Pasch T, Pietrokovski S, et al. A mesophilic cysteine-less split intein for protein trans-splicing applications under oxidizing conditions. *PNAS*;Available from: www.pnas.org/cgi/doi/10.1073/pnas.1909825116.

[25] Shi J, Muir TW. Development of a tandem protein trans-splicing system based on native and engineered split inteins. *Journal of the American Chemical Society*. 2005;127(17):6198–6206. Available from: https://doi.org/10.1021/ja042287w.

[26] AGScientific. *TCEP HCl vs DTT: Preferred Reducing Agents*;. Available from: https://bit.ly/2PdzlR8.

Computational Fluid Dynamics Modelling of a Wood-Fired Cookstove with a Forced Convection Air-Tray

Maximos Sifneos and Vinay Patel
Department of Chemical Engineering, Imperial College London, U.K.

Abstract
Following global impetus towards clean energy, a renewed focus has been placed on the design of clean and affordable biomass cookstoves which serve household needs of 2.4 billion people. In this work, a CFD model of a forced convection air-tray is developed and integrated with a heat transfer model of a Prakti wood-fired cookstove. Simulations were performed in ANSYS Fluent for a variety of fan positions and tray geometries to maximise the volumetric flowrate from the fan to the air-tray, enhance uniform air flow distribution through the tray outlets, and minimise energy losses. It was found that placing the air-tray inlet at the side of the tray and far from the outlets generated the highest inlet air velocity and improved air flow distribution. Curving the edges of the rectangular tray design yielded superior performance with air-flow increasing by 9.2%, outlet velocity uniformity by 2.7% and energy losses decreasing by 13.0%. Experimental validation was sought for the system's operating pressure predicted with CFD by using a pressure transducer attached to a DC voltmeter, with experimental readings being 33.6% lower than computational predictions. A heat transfer model of the integrated cookstove/air-tray design highlighted the benefits of forced convection in improving convective heat transfer, promoting upward air-flow in the stove, and reducing backflow at the cookstove inlet.

Keywords: *air-tray, biomass, computational fluid dynamics, cookstove, forced convection, natural convection, radiation*

Introduction

Air pollution is a significant contributor to the exacerbation of respiratory-related ailments including lung cancer, strokes and ischaemic heart illnesses. Non-developed countries are largely at risk from such effects due to the use of antiquated and highly polluting technology, with India a prime example, being one of the most highly polluted countries in the world (Balakrishnan , et al., 2019). A primary factor leading to air contamination in developing countries is the use of inefficient cookstoves for household tasks such as heating and cooking. More than 2.4 billion people globally currently rely on small biomass cookstoves for domestic use (IEA, 2017), accounting for c.15% of global energy use (Jeuland & Pattanayak, 2012). Inefficient burning of biomass releases substantial quantities of gaseous carcinogens including carbon monoxide (CO), nitrogen oxides (NO_x), and sulphur dioxides (SO_x) as well as solid particulate derivatives (Kumar, et al., 2013). The World Health Organisation (2018) concluded that close to 4 million people are succumbing to life-threatening illnesses due to high levels of domestic pollution as a direct result of inefficient cookstoves. In addition, cookstoves expel a variety of greenhouse gases (GHGs), and account for 2.3% of global CO_2 emissions (Lacey, et al., 2017).

Prakti is a social enterprise established in India which has partnered with Imperial College London to help combat indoor air pollution by producing cheap, clean and efficient biomass cookstoves for deprived communities. By finding engineering solutions to improve fuel burning, Prakti cookstoves offer environmental benefits in the form of fewer GHG and other harmful emissions, simultaneously providing economic savings to rural communities by reducing the fuel requirements of the stove; Prakti estimates an 80% reduction in fuel inputs and smoke generation in cleaner and more efficiently burning cookstoves (Prakti, 2019). Previous experimental work by Imperial College London found that the thermal efficiency of Prakti's cookstove design lies between 10-20% (John, 2019), a Tier 2 level design based on thermal efficiency as classified by the Voluntary Performance Targets laid out by the Clean Cooking Alliance (2019). The targeted reduction in emissions and fuel requirements would help to improve the design to Tier 5 level, the highest possible, which aligns with the WHO "guidelines for indoor air quality" (World Health Organisation, 2014).

Computational Fluid Dynamics (CFD) offers a time and resource efficient method to test the benefits of new designs in reducing emissions. Its advantages over experimental methods also include its adaptability to different designs, which ordinarily would have to be manufactured to be tested. However, CFD modelling often entails a large degree of approximations and can require substantial computational power. In addition, the process of wood burning in a cookstove is not a well understood or documented process and can be difficult to simulate. Modelling an unsteady, three-dimensional flame also poses significant challenges. Thus, computational modelling requires experimental validation to verify it is an adequate representation of the physicochemical processes it simulates.

It has been shown that combustion efficiency can be increased to improve thermal efficiency and to reduce harmful emissions by up to 90% in cookstoves utilising forced convection (Anenberg, 2012). This study aims to investigate the use of a forced convection air-tray for a Prakti wood-fired cookstove using CFD. More specifically, the position of the tray's air inlet and different tray shapes are examined to maximise the flowrate of air into the tray, achieve uniform outlet flow distribution and minimise energy losses in the air-tray. Experimental validation was sought for the operating point predicted by CFD. Additionally, this report seeks to assess the effectiveness of the air-tray to promote upward flow, reduce backflow and improve convective heat transfer in a heat transfer model of the cookstove.

Background

Over the past 30 years, there has been considerable effort in improving the design of biomass cookstoves after acknowledging their adverse effects on individuals' health and environmental pollution. Conversely, limited advances have been made in the use of computational fluid dynamics to improve biomass cookstoves; only 30 of the 500 journal articles associated to biomass cookstoves that were examined by MacCarty & Bryden (2016) utilised CFD.

In the 1980's, efforts were made by Bussman & Prasad (1986) and De Lepeleire & Christiaens (1983) to predict the fluid dynamics of flow and heat transfer using simple computational models. These models were soon developed and optimised at the end of the last decade by Gupta and Mittal (2010), where parameters such as spacing between the stove and pan, and primary and secondary air inlet size were investigated. Further development of these studies has been undertaken by Ali and Wei (2017), utilising ANSYS Fluent to analyse the effects of combustion on heat transfer, generating temperature gradients and velocity profiles within the stove. Nevertheless, there has been minimal experimental validation of the computational models (MacCarty & Bryden, 2016) to verify that they are an adequate representation of the actual processes they model.

A more specific study to the Prakti cookstove was completed by Hamilton & Lucas (2017) that couples flow and heat transfer using STAR-CCM+. In this report, a similar design to the current cookstove that Prakti manufactures was considered, with a natural convection air inlet at the front of the stove, as well as a design that promotes mixing in the top part of the chamber using a helical structure and secondary air inlets using natural convection placed at the height of the helix. The temperature of the woodblocks was fixed at 1000K and 1300K. Additionally, combustion models from Mapelli et. al. (2013) and radiation models (Discrete Ordinate Method) were adopted in their CFD simulations resulting in experimental and CFD discrepancies ranging from 38%-44%.

A study from Imperial College London by Wong & Ong (2019) built upon the findings of Hamilton & Lucas (2017) and focused on improving combustion modelling. The combustion model utilised was based on previous work by Malte (1996) on wood dust combustion kinetics. Further experimentation work was carried out in order to validate the CFD models, which were developed using ANSYS Fluent. The CFD modelling provided useful insights regarding cookstove emissions, which are directly relevant to Prakti, by considering solutions to allow for more air to enter from the bottom of the cookstove into the chamber to promote complete combustion. This was achieved by raising the wood blocks 4cm from the cookstove floor, resulting in CO emissions reduction by c.50%. Cookstove temperatures obtained using CFD were c.100-200K lower than experimental values.

John (2019) considered the design of a forced convection air-tray to improve burning efficiency in biomass cookstoves using CFD, a research area without much prior investigation. ANSYS Fluent was used to model various air-tray geometries including changing the orifice number, the diameter of the holes and their arrangement. Experimental validation was sought for the pressures and volumetric flows predicted by the model within the tray. It was found that a setup with 21 holes produced the highest volumetric flowrate from the fan. However, the air-tray was not integrated with the cookstove and its benefits were thus not investigated. In addition to the air-tray, the cookstove model developed by Wong & Ong (2017) was replicated. However, the wood material was changed to char and the combustion of char, previously unmodelled, was simulated as a volumetric heat source term on the woodblocks. Temperatures were over-predicted in the CFD model by 200-300K compared to experiments. Nevertheless, the distribution of temperature in the cookstove closely resembled experimental observations, with higher temperatures being recorded at the back of the stove. In addition, the CFD model predicted significant backflow at the primary inlet, considerably hindering combustion efficiency.

Another study by Husain et. al. (2019) sought to optimise the geometric design of a wood-fired cookstove to promote uniform distribution of air within the combustion zone. It investigated the use of a forced draft air inlet with staged introduction of air into the cookstove achieved via an annular space surrounding the stove. This design was found to improve the uniformity of airflow within the cookstove. However, the combustion was modelled homogeneously with gaseous fuel instead of considering the burning of the wood itself. Nevertheless, the report is a clear example of the advantages of CFD modelling in estimating the performance of a variety of designs, providing reasonably accurate insights into temperature, velocity and heat profiles within the cookstove, which would have been resource and time intensive to obtain using traditional experimental methods.

Methods

This study utilises ANSYS Fluent v19.0 R3 to build a CFD model of a forced convection air-tray and a woodfired cookstove. A simple experimental set-up is utilised to validate the operating point of an air-tray design.

Air-Tray & Cookstove: Geometry

Both the air-tray and the cookstove geometries were modelled using SpaceClaim. For the base case, the air-tray was modelled as a 180mm x 120mm x 10mm rectangular box. The air inlet was a 10mm x 8mm rectangle, and the outlets were sketched as 57, 2D-holes of 2mm diameter. An example of one of the designs can be seen in Figure 1a. Alternative designs with curved edges and internal baffles were built from the base case using the "Pull" tool to fillet the desired edges.

The modelled air-tray was placed beneath Prakti's cookstove design. The primary air inlet was placed at the front of the stove with dimensions 150mm x 100mm. The cylindrical, vertical section of the cookstove was modelled as a cuboid to obtain a more coherent mesh, and the cookpot was shaped as a flat sheet at the top of the vertical section. Air exits the stove through a 90mm x 27mm rectangular section at the side and near the top

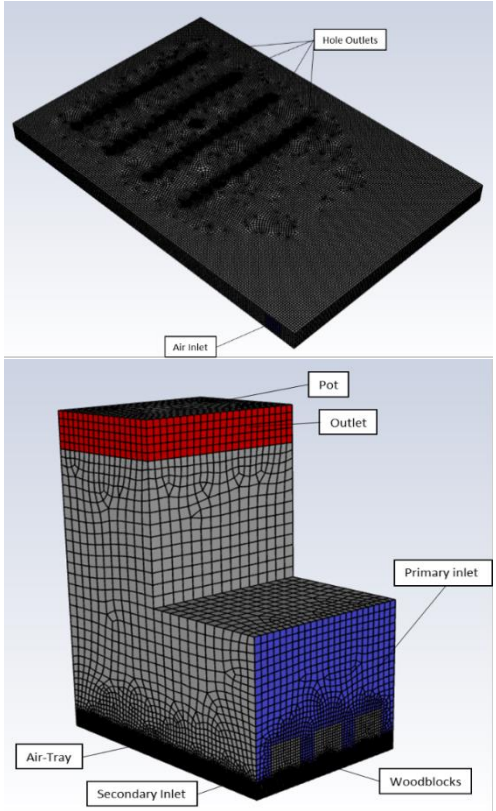


Figure 1a (Top): *Mesh used for CFD modelling of forced convection air-tray with 57 holes and rectangular side inlet.*
Figure 1b (Bottom): *Mesh used for CFD modelling of wood-fired cookstove with a forced convection air-tray. Adapted from Prakti's existing design.*

of the vertical section of the stove, as shown in Figure 1b. Three woodblocks (24mm x 24mm x 175mm) were placed at the base of the stove and covered 41 of the 57 air-tray outlet holes. To ensure coherent meshing, all bodies and faces were joined into a singular part.

Air-Tray & Cookstove: Mesh

Meshing of the geometries was performed in ANSYS Mesher. In order to minimise computational cost without sacrificing significant accuracy, a MultiZone, Hexa Core approach with localised face sizing was used to mesh rectangular shaped air-trays, resulting in meshes with c.300,000 elements. For non-rectangular trays where a Hexa-based approach is not suitable, a mesh was generated with tetrahedral components, leading to meshes with c.900,000 elements.

For the cookstove, a Hex Dominant, Quad/Tri meshing approach with individual face and body sizing resulting in 493,490 elements was found to be the best compromise between computational cost and accuracy.

Air Tray & Cookstove: Steady Solver Set-Up

The model was set-up and solved using ANSYS Fluent. A steady solver was used for most simulations due to the computational cost of using a transient solver, which was prohibitive given the project's timeframe, despite the modelled process being inherently transient (see *Air-Tray: Transient Solver Set-Up)*. A pressure-based solver was used as it is the most popular solver type for incompressible or mildly compressible flows (ANSYS, 2019) and was found to converge faster than a density-based solver. A SIMPLEC Pressure-Velocity coupling scheme was preferred over a Coupled option as it offers sufficient accuracy with lower memory requirements for reasonably uncomplicated, steady flows, and low mesh skewness (ANSYS, 2019). Internal cell gradients were determined using the Least Squared Cell Based method, which is less computationally intensive than the Nose Based approach (ANSYS, 2019). Finally, a second order discretisation scheme was employed to calculate flow and scalar terms in the governing equations.

Air-Tray: Transient Solver Set-Up

Transient simulations were performed to ascertain whether steady-state calculations adequately described the process simulated for the purposes of this investigation. A non-dimensional time was calculated as per Equation (1):

$$\tau = \frac{t}{T} \tag{1}$$

where t is the time taken to asymptotically reach a steady state value of static pressure and $T = L/U_{av}$; where L is characteristic length (thickness of plate) and U_{av} is the average steady state velocity at the inlet.

In order to change from a steady state to a transient simulation, the transient solver option was selected under the "General" tab in ANSYS. It was also necessary to define a time step per iteration whilst ensuring convergence for each step. For this, an adaptive Courant, Friedrichs and Lewy (CFL) based method was adopted, where a CFL number (or "Courant" number) of 1 was chosen. This is the recommended value for convection dominant flows (ANSYS, 2019). The Courant number is defined as:

$$C = \frac{\Delta t}{u/\Delta x} \tag{2}$$

where C is the non-dimensional Courant number, Δt is the time step per iteration, u is the average velocity in that cell and Δx is the width of the cell.

The denominator, a characteristic time scale, was calculated locally at each node and the appropriate time step was calculated by using Equation (2). This approach reduces the number of iterations required to achieve convergence per time step. An upper bound of 1000 iterations per time step was imposed to limit total computational time and ensure that each time step converged sufficiently.

Air Tray: Solver Set-Up: Models

To simulate convective transport, the mass and momentum conservation equations were solved. The gravitational term was de-activated due to the negligible height differential between the inlet and the outlet (less than 1cm). Due to small pressure gradients within the tray, heat and compressibility effects were neglected and the fluid was modelled as air at 15^0C.

Using a velocity of 10 $m.s^{-1}$, air density at 15^0C (1.225$kg.m^{-3}$), the thickness of the tray as the characteristic length (10mm) and air viscosity at 15^0C, a Reynolds number of 6625 was determined, which is above the internal flow turbulence transition threshold of 2300 (White, 2000). The RANS k-ε realisable method was selected to model turbulent fluctuations as it accurately predicts planar and round jet spreading rates. It was preferred over the standard k-ε model due to

superior performance for flows involving large streamline curvature and eddy vortices. Other turbulence models were considered, including the k-ω and Reynold Stress Model, which offer better near-wall performance, but weren't found to offer substantial benefits compared to the k-ε model. Finally, standard near-wall functions were used to solve the near wall regions, as the ε-transport equation contains a term which is not solvable at the wall (ANSYS, 2006).

The energy dissipated by the flow in the tray was estimated by calculating the difference between the static pressure and kinetic energy flux terms at the inlet and the outlet. The terms were determined by Fluent for each cell using a "Custom Field" function and integrated over the total inlet and outlet areas, according to Equation (3):

$$Energy\ Flux = \int [|u| \times (p + \tfrac{1}{2}\rho|u^2|)]dA \qquad (3)$$

where u is the cell velocity, p is the static pressure, ρ is the density of air and dA is a differential area.

Cookstove: Solver Set-Up: Models

All model settings active for the air-tray were retained in modelling the cookstove. The stove's operation was approximated as a pure convective and radiative heat transfer model; the combustion of wood volatiles in the gas phase and the formation of a flame were not considered in this study.

Due to convergence difficulties with the realisable k-ε model, the standard k-ε model was used.

Additionally, convective heat transfer was simulated by enabling the governing energy equation. The wood blocks were modelled as char to simulate the surface combustion of char following the devolatilization of wood volatiles. As the combustion process is not well understood and difficult to model it was approximated by setting the charred wood as volumetric heat sources at 1674kW.m^{-3}. The source term was determined iteratively to give an approximate char temperature of 1000K (Evans & Emmons, 1977). Moreover, gravitational acceleration was set at 9.81m.s^{-2} and the density of air was modelled to vary according to the compressible ideal gas equation of state. The specific heat capacity and thermal conductivity correlations for char and air are summarised in Table 1.

Table 1: Thermal properties of char and air. Air properties taken from Felder & Rousseau (2001), thermal conductivity of char from Brown (1972) and heat capacity from Gupta et. al (2003).

Species	Specific Heat Capacity (J.kg^{-1}.K^{1})	Thermal Conductivity (W.m^{-1}.K^{-1})
Char	$440 + 2T - 5.96 \times 10^{-4}T^2$	0.221
Air	$0.028 + 0.197 \times 10^{-5} \times T + 0.48 \times 10^{-8} \times T^2$	$0.043 + 8 \times 10^{-5}T - 1 \times 10^{-8}T^2$

Radiation was modelled using the Discrete Ordinates (DO) Model. The DO model converts the radiative transfer equation into a transport expression solving for radiative intensity in the cartesian coordinate system. It is the recommended model for localised sources of heat. Although the P1 model is less computationally expensive, it usually over-estimates radiative fluxes. A gray radiation model was assumed, with the woodblocks modelled as perfect emitters, the pot as a perfect absorber and the cookstove walls as perfect reflectors.

Air Tray: Boundary Conditions

To simulate forced air convection from the fan, the inlet was set as a pressure or mass inlet boundary condition, the value of which depended on the operating point for each system (see *System Operating Point Determination*). Atmospheric pressure conditions were imposed at the outlets. Initial guesses for the turbulent intensity and viscosity ratio of 5% and 10, respectively, were enforced at both the inlet and outlet. A no-slip condition was imposed on the walls.

Cookstove: Boundary Conditions

The primary, natural convection, air inlet at the front of the stove was added as a pressure boundary at atmospheric conditions and the air-tray holes were converted to an interior part. All boundary conditions are summarised in Table 2.

Table 2 - Summary of Cookstove Boundary Conditions

Boundary	Property	Value
Natural Convection Air Inlet	Gauge Pressure Inlet (Pa)	0
	Temperature (K)	299
	Turbulent Intensity (%)	5
	Viscosity Ratio	10
Forced Convection Air Inlet	Gauge Pressure/Mass Inlet	Operating
	Temperature (K)	299
	Turbulent Intensity (%)	5
	Viscosity Ratio	10
Air Outlet	Gauge Static Pressure Outlet (Pa)	0
Woodblocks (Wall)	Thermal Conditions	Coupled
	Material	Char
	Internal Emissivity	0.96
Pot (Wall)	Heat Transfer Coefficient (W.m^{-2}.K^{-1})	36.9
	Free Stream Temperature (K)	353
	External Emissivity	0.7
	External Radiation Temperature (K)	299
	Internal Emissivity	0.4
	Shell Conduction Material	Al
	Shell Conduction Thickness (m)	0.002
Stove (Wall)	Heat Transfer Coefficient (W.m^{-2}.K^{-1})	26.2
	Free Stream Temperature (K)	353
	External Emissivity	0.7
	External Radiation Temperature (K)	299
	Internal Emissivity	0.4
	Shell Conduction Material	Steel
	Shell Conduction Thickness (m)	7

System Operating Point Determination

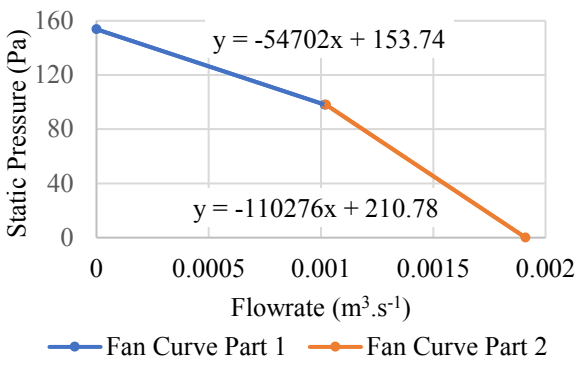


Figure 2: Simplification of ARX CeraDyna Centrifugal fan performance curve into 2 straight line segments. The high pressure, inefficient region is represented by the top left (blue) line, and the high efficiency region by the bottom right (orange) line. Although not reflected in the simplified fan curve, the stall region of the fan lies near the intersection of the line segments.

The fan performance curve supplied by the manufacturer, ARX, of the CeraDyna Centrifugal fan was simplified into two line segments representing the high and low efficiency regions, as seen in Figure 2.

Several numerical simulations were run in Fluent using a variety of inlet conditions to obtain data points on the system resistance curves of each geometry. A pressure inlet boundary condition was imposed in each case due to superior convergence performance over mass or velocity boundaries. From the momentum conservation equation at constant elevation, it is known that static pressure is proportional to the square of the volumetric flowrate, and the data points were thereby fitted using a quadratic fit. The system operating point was determined from the intersection of the fan and system curves. The estimated operating point was subsequently inputted as a mass inlet boundary condition in Fluent to verify it lied on the system curve.

Experimental Methodology & Set-Up

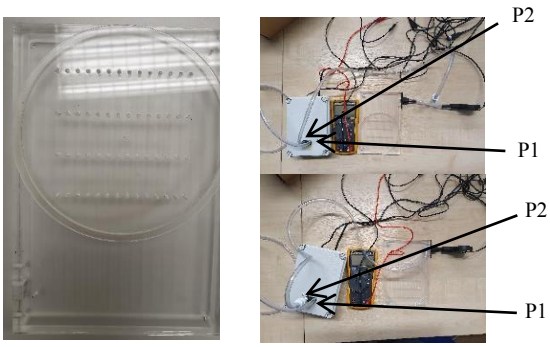


Figure 3a (left): *Perspex air-tray with drilled array of holes and an inlet on the bottom left of the tray where a fan is connected.*
Figure 3b (right): *2 alternative set-ups of the experiment. The original design (design 1, top right) and a modified design (design 2, bottom right) to minimise fitting losses.*

A Perspex plastic tray was used to replicate an air-tray design as shown in Figure 3a. The geometry was manufactured by drilling 57 holes into the tray to replicate the outlet holes. The drilled holes were uniformly distributed in a rectangular 4x14 array, with a diameter of 2mm with an additional hole in the centre of the array. A rectangular opening was cut on the bottom left of the tray to allow flow of air from the centrifugal fan to enter the air-tray. The fan was attached to the tray via a series of plastic tubes, serrated fittings and a T-junction which allowed measurements of static pressure to be taken during operation. Figure 3b represents two ways in which the fan was connected to the air-tray with the objective to minimise losses energy losses from fittings and other flow restrictions.

To measure the static pressure generated by the fan at operating conditions, a pressure transducer was connected to a highly sensitive DC voltmeter, as shown in Figure 3b. The difference in static pressure between the fan and the atmosphere was reflected in a voltage reading across the P1 and P2 transducer ports. The P1 port was connected to a tube exposed to ambient air, whilst the P2 port was fixed with a plastic tube linked to the point at which the pressure measurement was desired. The position at which the pressure was measured differed for each design; for design 1 (Figure 3b, top) the pressure was measured using a T-junction near the fan outlet. Due to the large number of fittings providing resistance to flow, a second set-up was tested (Figure 3b, bottom), in which the pressure was measured at a 2mm drilled hole just after the tray inlet. The plastic tube leading the P2 port was sealed with BlueTac to minimise air flow loss outside the closed system. To convert the voltage measurement to a pressure difference, a Pressure-Voltage calibration curve prepared by John (2019) for an identical set-up was used. However, the atmospheric voltage reading was adjusted to account for different atmospheric conditions between the measurements taken by John (2019) and in this investigation, assuming the slope of the straight line remained constant. Equation (4) below shows the adjusted calibration equation:

$$P = 171V - 900.26 \tag{4}$$

where P is the static pressure in *Pa* and V is the voltage (V) measured across the transducer ports P1 and P2.

In both cases the voltmeter was allowed to settle for 2 minutes before taking the readings, and a total of 5 readings were taken per set-up.

Results & Discussion

ANSYS Fluent Results: Transient Solver

At first, both steady and transient simulations were performed to ascertain whether steady state modelling was sufficiently accurate to describe the system for the purposes of this investigation. This was pertinent to this study, as the computational cost of using a transient solver for all simulations would have been prohibitive.

Adopting the methodology from *ANSYS Fluent: Transient Solver Setup*, the dimensionless time, τ, was determined for both the front inlet (Figure 5c) and side inlet (Figure 5b) geometries. The results have been summarised in Table 3.

The dimensionless times were found to be significantly larger than unity which implies that the times taken to reach the steady state static pressure values are much larger than the respective characteristic times and thus, the flow has significantly developed when compared to the flow at time, T. This indicates that

for means of comparison, a steady state simulation is adequate to describe the process.

Table 3: Summary of the dimensionless time calculations. The characteristic time scales, *T,* the characteristic length scale was the thickness of the plate, *L*, whilst the inlet average velocity, U_{av}, was determined from the steady state runs. The time, *t*, was the time taken to asymptotically approach the steady state static pressure for the respective geometry.

	Front Inlet	**Side Inlet**
t (s)	0.082	0.125
L (m)	0.010	0.010
U_{av} (ms^{-1})	11.8	16.6
$T = L/U_{av}$ (s)	0.001	0.001
τ (-)	97	206

Figure 4 depicts velocity contours of the two geometries examined after 8.2ms for the front inlet and 12.5ms for the side inlet. A comparison of these contours with the steady state vector plots for the same configurations in Figures 5a & c, shows that the flow patterns predicted are similar in their core features (general shape and order of magnitude). This indicates that the flow is adequately developed after a sufficiently short time making it justifiable to use a steady state approximation in modelling flow in the air-trays.

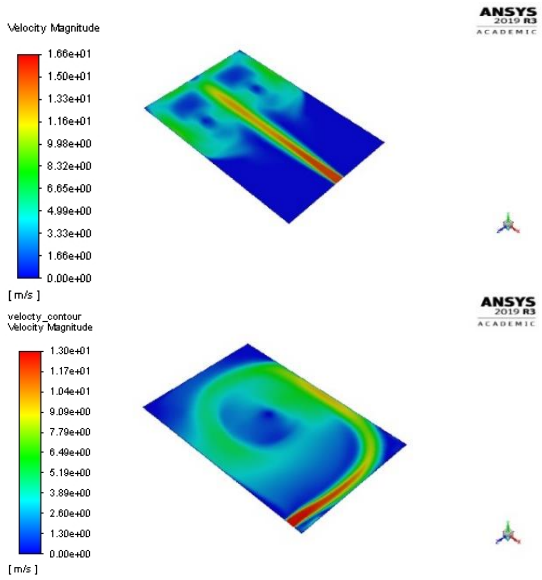


Figure 4: Velocity contour plots of both the front inlet (top) and side inlet (bottom) geometries after running for 8.2ms and 12.5ms, respectively.

ANSYS Fluent Results: Air-Tray Operating Point

Four different positions for the tray's air inlet were examined as illustrated in Figure 4; at the side of the tray, at a distance from the holes (1), at the front of the tray, away from the holes (2), at the side of the tray in line with the centre of the hole array (3), and at the back of the tray near the holes (opposite inlet (2))(4).

The system resistance curves for the four geometries were plotted in Figure 5a. Geometry 1 and 2 (grey and yellow curves) exhibit the highest inlet volumetric flowrates at $1.19\times10^{-3}m^3.s^{-1}$ and a corresponding static pressure of 79.5Pa. This highlights the benefits of placing the secondary inlet at a distance from the tray's outlet holes, thereby minimising the resistance to the incoming air caused by a high-pressure region created near the hole outlets. A clear selection cannot be made

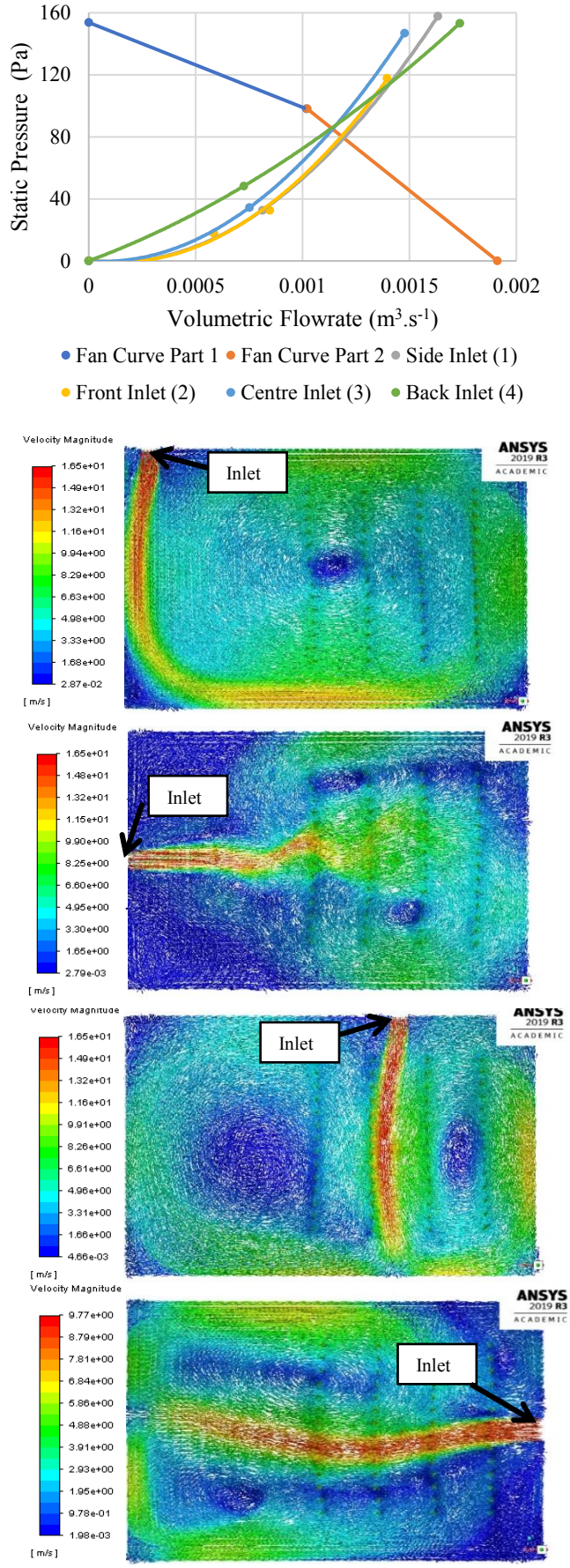


Figure 5: Different positions for the air inlet in the forced convection air-tray. **Figure 5a.** system curves for different geometries overlaid with fan performance curve; Geometry 1 in grey, geometry 2 in yellow, geometry 3 in green, geometry 4 in light blue; **Figure 5b**. Velocity vector plot for geometry 1; **Figure 5c.** Velocity vector plot for geometry 2 **Figure 5d**. Velocity vector plot for geometry 3; **Figure 5e.** Velocity vector plot for geometry 4.

between geometries 1 and 2 based on the magnitude of the inlet volumetric flowrate of air. The uncertainties incurred as a result of numerical method imprecisions from Fluent, the inaccuracy of the system curve quadratic fit, and the error introduced by linearizing the fan curve outweigh the marginal differences between the two geometries' operating points.

The vector plots in Figures 5(b)-(e) show the flow path and velocity magnitude of air in the tray due to the pressure differential between the fan outlet and the hole outlets. Geometry 1 demonstrates a more uniform distribution of air through the outlet holes compared to the other designs. It also shows fewer velocity 'dead-zones', as illustrated by the dark blue regions on the plots, especially in the near-hole regions. The swirling nature of the flow does however lead to holes on the perimeter of the array being favoured over those in the centre of the tray. Geometries 3 and 4 both have high velocity jets passing through outlet holes leading to highly non-uniform flow. In Geometry 2, the positioning of the inlet away from the walls of the tray does not promote mixing, thereby favouring those holes in the centre of the array, near the inlet. These observations are corroborated by the outlet velocity uniformity index, γ_a, which is a metric that measures how a field variable varies over a surface and is given by Equation (5):

$$\gamma_a = 1 - \frac{\sum_{i=1}^{n}[(|u_i - \overline{u_a}|)A_i]}{2|\overline{u_a}|\sum_{i=1}^{n} A_i} \quad (5)$$

where u_i being the cell velocity, A_i the area of the cell, and u_a the average velocity over all cells.

The index takes values between 0 and 1, and the higher the value, the more uniform the flow. Geometry 1 was found to have the largest uniformity index at 0.914 followed by geometries 2 and 3 with indices of 0.905 and 0.902 respectively. Geometry 4 displayed the lowest index due to its proximity to the orifices and large distance from the tray walls. One shortcoming of the uniformity index is that it also considers differences in the velocity of air between individual cells making up each outlet. Nevertheless, given that the results validate the observations made from the velocity vector plots, they were deemed to be a fair reflection of the velocity distribution.

Table 4 summarises the energy dissipation of the air stream in the tray. Geometries 2 and 4, where the inlets direct the air-flow towards the centre of the tray, experience lower energy loss between the inlet and the outlet compared to those with the air inlet positioned on side of the tray. This is likely due to the energy loss caused by the inelastic collision of the inlet jet with the opposing wall.

Table 4: Energy loss of air stream in the air-tray between the inlet and the outlet.

Geometry	Energy Dissipated (%)
1	77.1
2	67.4
3	74.4
4	66.9

Experimental Results: Air-Tray Operating Point

The results from the experimental set-up are summarised in Table 5. Although a discrepancy between the operating static pressure obtained using CFD and the experimental procedure of more than 33% was obtained for both configurations, the error was smaller using the modified design (b).

Table 5: Experimental results of the two designs as shown in Figure 3a & b. Mean static pressure reading are converted to static pressures. An absolute percentage error between the simulation value and the set-up value was also calculated to represent the magnitude of losses present.

Set-Up	Voltage (V)	Static Pressure (Pa)	Discrepancy (%)
Original (a)	5.495	38.985	51.0
Modified (b)	5.57	51.81	33.6

For the original set-up (a), there were significant energy losses in the T-junction due to the large number of fittings, which caused a reduction in the velocity entering the air-tray, resulting in lower than expected pressure readings. In modified set-up (b), the accuracy of the measurements was improved by reducing the number of fittings and placing the fan closer to the tray inlet. However, energy was lost from the system due to the imperfect sealing of the plastic tubing connected from P2 to the outlet hole near the inlet. The use of silicone grease, similar to that used by John (2019), instead of BlueTac, would serve to reduce energy losses due to inadequate sealing further. Other sources of discrepancy between experimental and computational predictions arise from the pressure measurement not being taken directly at the inlet of the tray, as is the case with the computational model. In the first design there are several fittings after the location of the pressure measurement up to the inlet, and in the second set-up the pressure measurement is taken at a small hole after the inlet.

ANSYS Fluent Results: Alternative Designs

Three alternative designs were created and tested in ANSYS Fluent, building upon the findings in *ANSYS Fluent Results: Air-Tray Operating Point* and tries to further improve air-tray performance. All designs described in this section use design 1, with the inlet at the side of the tray and away from the outlet holes, as the base case. In the first alternative design (5) (Figure 6a), the 3 corners of the air-tray away from the inlet were curved, with a fillet radius of 10mm. The second configuration (6) (Figure 6b) was similar to geometry 5 but with a fillet radius of 86mm on the corner opposite the inlet, in order to streamline the curvature of the flow path, minimise energy losses and guide the flow to the centre of the hole array. For geometry 7 (Figure 6c), in addition to the increased curvature of design 6, two additional baffles, with a 51mm arc radius, were added either side of the outlet holes.

Design 5 showed superior performance with regards to inlet flow rate compared to Design 6, with a flow rate of 1.30×10^{-3}m.s^{-1} compared to 1.24×10^{-3}m.s^{-1}. The first design also marginally outperformed the second alternative in terms of flow uniformity, with outlet velocity uniformity indices of 0.939 and 0.935, respectively, as the larger filleting of the latter guides the flow to the rear of the tray more so than the former.

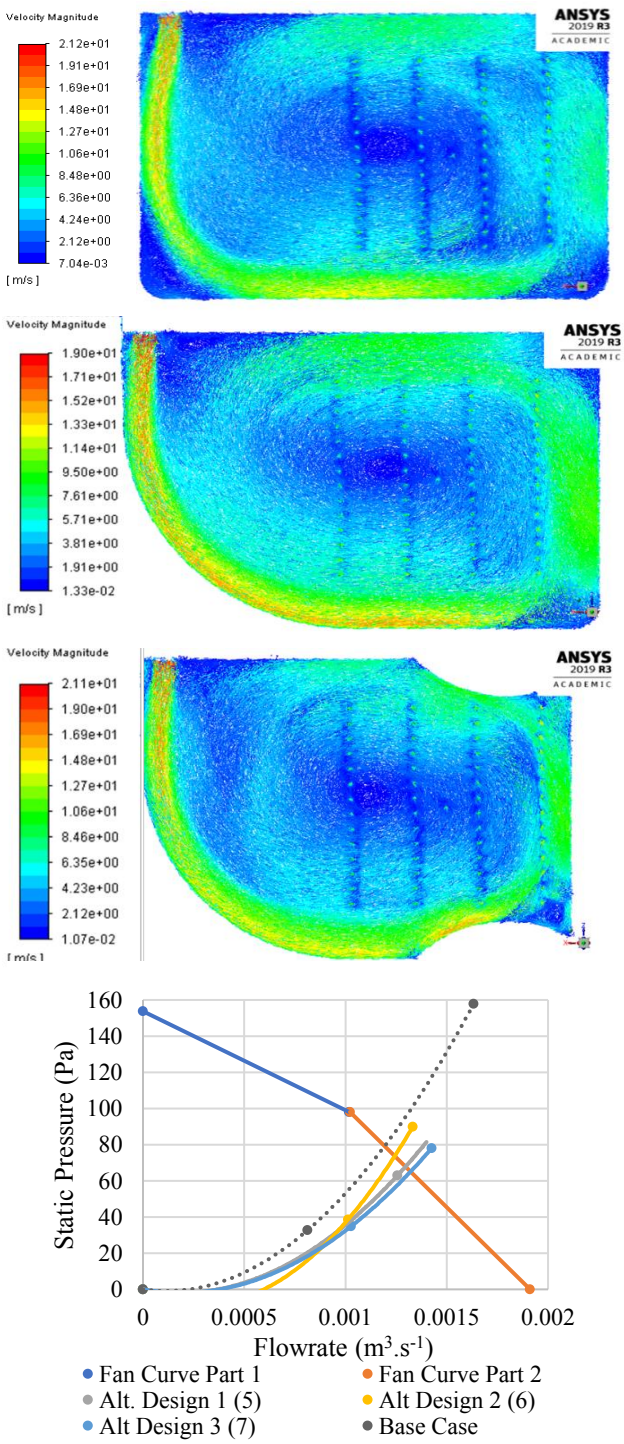


Figure 6: Alternative designs for the air-tray. **Figure 6a** is alternative design 1; **Figure 6b** is alternative design 2; **Figure 6c** is alternative design 3; **Figure 6d** is operating points of each design (in solid colours) with the original rectangular (base case) design (dotted grey line) added for comparison.

The third alternative design was intended to push air-flow towards the centre of the tray in order to avoid air preferentially exiting the tray through the holes at the rear and achieve a greater flow uniformity from the outlets, as well as maximising flow velocity. However, there was a reduction in the uniformity index compared to the other alternative geometries, at 0.927, despite a small increase in inlet volumetric flowrate at 1.32×10^{-3}m.s^{-1}. The reason for this reduction can again be attributed to the flow being pushed towards the back of the air-tray and preferentially out the rear holes. The intended aim of a more uniform distribution of air and a reduction of flow build up near the back of the plate was, however, not achieved. Nevertheless, the baffles do succeed in making the flowlines curve earlier than in designs 5 and 6. Thus, placing the baffles closer to the middle of the tray length, rather than near the rear, could achieve the design objective.

Nevertheless, all three of the alternative designs proved to be an improvement from the rectangular base case design, as seem from the system curves in Figure 6d. Both an increase in volumetric flow and a reduction in static pressure at the inlet was observed, accompanied with more uniform flow. In addition, though energy loses in the three alternative designs were comparable, the curved designs showed a c.10% reduction in energy losses compared to the rectangular design. The ease and simplicity of implementing curved edges to the existing rectangular design make this an attractive solution to improve the performance of the air-tray from both an engineering and a manufacturing perspective. Finally, given the improved performance of both designs 5 and 6 compared to design 1, and superior performance of 5 compared to 6, there is likely an optimum filleting level for the edge opposite the air inlet which matches the natural flow path of the incoming air stream and maximises the volumetric flow delivered by the fan to the tray.

ANSYS Fluent Results: Cookstove and Air-tray

Figure 7a depicts temperature contours of a vertical slice of the cookstove generated using CFD and illustrates the temperature distribution in the stove. A high temperature region of c.1000K is found along the length of the woodblocks. There is a significant temperature decrease between the woodblocks and the surrounding air, reflecting air's poor heat conducting ability. As predicted in previous work by Imperial College London (Ong & Wong, 2019; John, 2019), the temperature of air is higher towards the back of the stove and is a result of the direction of natural convective flow from the primary air inlet which promotes convective heat transfer at the back of the stove. Additionally, the cookstove's thermal efficiency was 7.5%; although this notably lower than a Tier 4 or Tier 5 level (>40%), it is expected to increase substantially in future work when, for example, the combustion of wood volatiles is included. A small region of high temperature is maintained away from the cookstove walls (indicated by the arrow in Figure 7a) due to the high velocity, hot air emanating from the forced convection air-tray, which was preheated by the woodblocks before entering the cookstove. This highlights the potential of forced convection air-trays in achieving more uniform air flow and temperature distribution in the cookstove, thereby improving the burning efficiency of the wood volatiles and resulting in a cleaner and more energy efficient cookstove. However, improvements are required in the design of the air-tray to improve air mixing inside the cookstove, which is insufficient in the current design, as illustrated in Figure 7c. One method proposed by Ong & Wong (2017) would be to create additional supports in the cookstove to lift the woodblocks above the air-tray

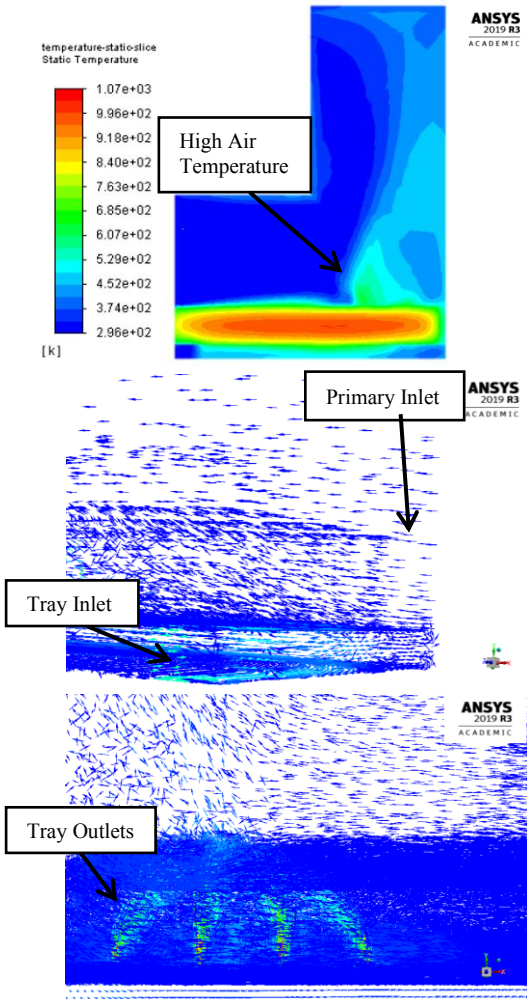


Figure 7: Integrated air-tray and cookstove; **Figure 7a:** Temperature contours on vertical cross-section of cookstove side view**; Figure 7b:** Velocity vectors at cookstove and air-tray inlets. Light blue vectors indicate higher velocity at the inlet of the tray. **Figure 7c:** Velocity vectors, side view, depicting high velocity air being ejected by the air-tray outlet holes to the cookstove interior.

outlets. This would result in marginal losses of air-flow velocity but substantial benefits in its flow rate and distribution into the cookstove.

Previous work by John (2019) on Prakti cookstoves without a forced convection air-tray indicated significant air backflow at the primary inlet. As illustrated by Figure 7b, no backflow was observed in the integrated air-tray and cookstove design. This is vital in ensuring complete burning of the fuel in the cookstove chamber thereby decreasing harmful emissions and improving thermal efficiency.

Conclusions & Recommendations for Future Work

The main objectives of this investigation were the modelling and optimisation of a forced convection air-tray for use as a secondary air inlet in a Prakti wood-fired cookstove, and the development of a heat transfer model of the cookstove with the tray. These objectives have been met. However, experimental validation for the CFD model of the air-tray was not achieved.

The transient CFD model justified the use of steady state simulations as an adequate approximation of the process. The dimensionless times for the front and side inlets were 97 and 206 and indicated that the flow had developed sufficiently before reaching steady state.

Using a steady state solver, the best position for the tray's air inlet was far from the holes, with a volumetric flow rate of 1.19×10^{-3} $m^3.s^{-1}$, as this improves velocity distribution and increases the flow into tray.

Experimental work was also performed to try and validate the static pressure from the CFD simulations with a discrepancy of at least 33% largely due to energy losses around the fittings and the position at which the pressure was measured. Though progress has been made to match computational predictions to experimental data, it is still necessary to improve the experimental design before deciding on the accuracy of the CFD model. Progress can be made in reducing experimental error by improving the sealing of the air flow and reducing the number of fittings required in the set-up. It would also be beneficial to experimentally verify the air velocities predicted by the CFD model by using a large bubble column and measuring the time it takes for an air bubble from the fan takes to reach the top of the column.

Additional designs, of the air-tray were examined, with Alternative Design 1 having a higher inlet volumetric flowrate compared to Design 2. Although Design 3 had a higher inlet flowrate, the uniformity index was lower due to the air being preferentially distributed to the back of the tray. The curved designs all outperformed the rectangular design and the modifications are easily implementable from a manufacturing perspective. Moving forward, machine learning techniques can be utilised to efficiently test the effectiveness of new air-tray designs to promote cleaner burning by optimising the arrangement of the tray holes, the number of holes and the geometry of the air-tray. In addition, angled nozzles could be used to direct air away from the back of the stove towards the centre, to promote uniform distribution of air flow in the stove and support complete combustion. Moreover, designing an air-tray as a coiled tube with holes of increasing diameter in order to maximise the total flow of air from the tray into the stove could also be explored.

Integrating the forced convection air-tray with the cookstove proved to have significant benefits in promoting upward airflow and improving heat transfer. Although a large temperature drop was observed between the woodblocks and the surrounding air, a high air temperature region above the woodblocks was detected, the formation of which was aided by the air emanating from the air-tray below. Finally, the integrated cookstove-air-tray model predicted little backflow at the primary inlet of the stove, an improvement over previous designs without forced convection. The CFD model can be developed further with the addition of combustion, soot, and SO_x and NO_x models to the existing heat transfer model as well as validating the findings experimentally, by using a laminar flow hood, and specialist flow and emission sensors. The design of the geometry could also be improved by modelling the vertical section of the stove as a cylinder, rather than a rectangle.

Acknowledgements

We would like to thank everyone that helped us in this investigation: Dr Chris J. Tighe, Thomas Abadie and Prof. Omar K. Matar.

References

Ali, H. & Wei, T. (2017) CFD Study of an Improved Biomass Cookstove with Reduced Emission and Improved Heat Transfer Characteristics. *Journal of Clean Energy Technology.* 5(6), 427-432.

Anenberg, S. (2012) Clean stoves benefit climate and health. *Nature.* 490(7420) 343. Available from: doi: 10.1038/490343e.

ANSYS. (2006) *Modeling Turbulent.* Available from: http://www.southampton.ac.uk/~nwb/lectures/GoodPracticeCFD/Articles/Turbulence_Notes_Fluent-v6.3.06.pdf [Accessed 13 November 2019].

ANSYS. (2019) *ANSYS Help V19.0 R3.*

Balakrishnan , K. et al. (2019) The impact of air pollution on deaths, disease burden, and life expectancy across the states of India: the Global Burden of Disease Study 2017. *The Lancet Planetary Health.* 3(1), e26-e29. Available from: doi: 10.1016/S2542-5196(18)30261-4.

Brown, L. E. (1972) *An Experimental and Analytical Study of Wood Pyrolysis*. PhD Dissertation. Department of Chemical Engineering, The University of Oklahoma.

Bussman, P. & Prasad, K. (1986) Parameter Analysis of a Simple Woodburning Cookstove. Available from: https://www.semanticscholar.org/paper/PARAMETER-ANALYSIS-OF-A-SIMPLE-WOODBURNING-Bussmann-Prasad/7eecce3a765cf544f0cef04013738b9e1443ca2e [Accessed 9th December 2019].

Clean Cooking Alliance. (2019) *Voluntary Performance Targets.* Available from: https://www.cleancookingalliance.org/technology-and-fuels/standards/iwa-tiers-of-performance.html [Accessed 7th December 2019].

De Lepeleire, G. & Christiaens, M. (1983) Heat transfer and cooking woodstove modelling. *Proceedings of the Indian Academy of Sciences Section C: Engineering Sciences.* 6(1), 35-46.

Evans, D. D. & Emmons, H. W. (1977) Combustion of Wood Charcoal. *Fire Safety Journal.* 1(1), 57-66. Available from: doi: 10.1016/0379-7112(77)90008-X.

Felder, R. M. & Rousseau, R. W. (2001) *Elementary principles of chemical processes*, 3rd edition, John Wiley & Sons.

Gupta, M., Yang, J. & Roy, C. (2003) Specific heat and thermal conductivity of softwood bark and softwood char particles. *Fuel.* 82(8), 919-927. Available from: doi: 10.1016/S0016-2361(02)00398-8.

Gupta, R. & Mittal, N., 2010. Fluid Flow and Heat Transfer in a Single-Pan Wood Stove. *Int. Journal of Engineering Science and Technology,* 2(9), 4312-4324. Available from: https://www.researchgate.net/publication/50346354_Fluid_flow_and_heat_transfer_in_a_single-pan_wood_stove [Accessed 9th December 2019].

Hamilton, S. & Lucas, E. (2017) *Development of an Improved Biomass Cookstove Computational Fluid Dynamics Model with Experimental Validation.* Research Project. Department of Chemical Engineering, Imperial College London.

Husain, Z., Tiwari, S., Pandit, A. & Joshi, J. (2019) Computational Fluid Dynamics Study of Biomass Cook Stove - Part 1: Hydrodynamics and Homogenous Combustion. *Industrial and Engineering Chemistry Research.* Available from: doi: 10.1021/acs.iecr.9b03181.

IEA. (2017) *Energy Access Outlook 2017.* Paris: IEA. Available from: https://webstore.iea.org/download/summary/274?fileName=English-Energy-Access-Outlook-2017-ES.pdf [Accessed 9th December 2019].

Jeuland, M. A. & Pattanayak, S. K. (2012) Benefits and Costs of Improved Cookstoves: Assessing the Implications of Variability in Health, Forest and Climate Impacts. *PLOS One,* 7(2), 1-15. Available from: doi:10.1371/journal.pone.0030338.

John, S. (2019) *Development of a CFD Model for the Design of a Wood-Fired Cookstove with Forced Convection.* Master's Thesis. Imperial College London.

Kumar, M., Kumar, S. & Tyagi, S. (2013) Design, development and technological advancement in the biomass cookstoves: A review. *Renewable and Sustainable Energy Reviews,* 26(1), 265-285. Available from: doi:10.1016/j.rser.2013.05.010.

Lacey, F. G., Henze, D. K., Lee, C. J., van Donkelaar, A., Martin, R. V. (2017) Transient climate and ambient health impacts due to national solid fuel cookstove emissions. *PNAS,* 114(6), 1269-1274 Available from: doi: 10.1073/pnas.1612430114

MacCarty, N. & Bryden, K. (2016) Modeling of household biomass cookstoves: A review. *Energy for sustainable development,* 26(1), 1-13. Available from: doi: 10.1016/j.esd.2015.02.001.

Malte, P. C. (1996) *Development of a model for predicting the NOx emissions of burners fired with sawdusts and sanderdusts high in nitrogen.* Department of Mechanical Engineering, University of Washington.

Mapelli, F., Mereu, R., Barbieri, J. & Colombo, E. (2013) *CFD Feasibility analysis of an Improved Cook Stove (ICS) for electricity production.* Available from: http://hdl.handle.net/11311/738772 [Accessed 9th December 2019].

Ong, L. & Wong, C. (2019) Development of a Computational Fluid Dynamics Model with Experimental Validation of an Improved Biomass Cookstove. *ChE Research Imperial College London,* 1(1), 1-449.

Prakti. (2019) *About Prakti.* Available at: http://prakti.in/[Accessed 19th November 2019].

World Health Organisation. (2014) *WHO guidelines for indoor air quality: household fuel combustion.* Available from: https://www.who.int/airpollution/guidelines/household-fuel-combustion/IAQ_HHFC_guidelines.pdf [Accessed 9th December 2019].

World Health Organisation.(2018) *Household air pollution and health.* Available from: https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health [Accessed 19th November 2019].

White, F.M. (2000) Fluid Mechanics, 4th Edition. McGraw-Hill Publishing.

Performance of Modified Titanate Nanotubes in Adsorptive Desulfurisation Processes

Ben Laws, Amy Mabonga
Department of Chemical Engineering, Imperial College London

Abstract
Modern regulations place stringent limits on the permissible sulfur content of fuels that are difficult to meet efficiently with the methods currently available. Adsorptive Desulfurisation represents an innovative technology that could facilitate the effective production of Ultra Low Sulfur fuels when utilised in combination with Catalytic Hydrodesulfurisation. This study has therefore considered the application of titanate nanotubes as adsorbents in the removal of refractory sulfur compounds, namely dibenzothiophene and dialkyl-substituted dibenzothiophene, from samples of model diesel. Changes to equilibrium capacity and adsorbate selectivity were evaluated as adsorbent moisture content, operating temperature, adsorption time, and surface ion composition were varied, producing a superior understanding of the key performance drivers. At the concentration range evaluated, Freundlich isotherm behaviour was primarily exhibited. It was found that Hydrogen-form nanotubes offered a capacity of 9.22mgS/g when dried for 18 hours and tested at 20°C. Ion exchange with copper ions resulted in a capacity of 9.27mgS/g, which is predicted to improve substantially following synthesis refinements. While selectivity between thiophenic compounds is not competitive, steric hindrance and surface ion composition were shown to have a moderate impact in favour of dibenzothiophene. It was concluded that π-complexation and σ S-M adsorption mechanisms are likely to explain the adsorptive behaviour seen in this study, although DFT calculations are now needed to validate this hypothesis.

Keywords: SARS, Titanium, Nanostructures, Ion Exchange, Capacity, Selectivity

I. Introduction

Following the discovery of the full environmental and human health impacts of sulfur-based emissions in recent decades, the removal of sulfurous compounds from transportation fuels has become a key priority for both governments and refiners. These findings coincide with increases in the average sulfur content of extracted crude oil from 1.09wt% to 1.40wt% between 1990 and 2018 and growing demand for diesel in developing nations [1]. When combined, these factros highlight an urgent need for enhanced global desulfurisation capacity. A worldwide response culminated in the imposition of stringent regulatory limits on the sulfur content of diesel fuels, reducing permissible limits in many countries from 500ppm to 15ppm in 2007 [2].

While conventional techniques are adequate for the bulk removal of sulfur compounds, these methods have proven ineffective for the removal of aromatic thiophenes that is critical to achieving the deep desulfurisation now required. Adsorptive Desulfurisation is a developmental technology that, when employed in combination with traditional techniques, could offer the sustainable and highly efficient ultra-deep desulfurisation desperately needed in commercial refineries.

There has been extensive research activity in the field of two-stage desulfurisation processes, which incorporate traditional Hydrodesulfurisation (HDS) with next-generation techniques. The supplementary unit operation enables intense desulfurisation by specficially targeting unreactive sulfurous aromatics, thus facilitating a reduction in both the economic cost and environmental impact of the subsequent HDS stage. Adsorptive Desulfurisation (ADS) – already investigated in a number of pilot studies – has been shown to provide notable desulfurisation of refractory compounds at moderate operating conditions. This study considered the adsorptive performance of a process that utilises Titanate Nanotubes (TiNTs) as the adsorbent. This material has previously shown competitive adsorption capabilities at ambient conditions – a significant advantage, especially when combined with their highly scalable production process [3].

An extended investigation was undertaken to evaluate three objectives: *i)* the adsorptive performance and selectivity of TiNTs, *ii)* the factors that can enhance capacity, and *iii)* the impact of transition metal ion exchange on both the equilirbium capacity and the observed thiophene selectivity. This performance will be benchmarked against adsorption on comparative structures to verify whether TiNTs are a viable adsorbent for use in innovative desulfurisation procedures.

II. Technical Background

Catalytic Hydrodesulfurisation (HDS) is the conventional method for the large-scale removal of sulfur compounds from fractionated crude oil. Sulfur content is extracted following the reduction of target molecules, which is induced by a constant flow of hydrogen, exposure to a metal-doped zeolitic catalyst such as $CoMo/Al_2O_3$, and high temperature and pressures. To reduce sulfur concentrations below 500ppm as demanded by 21st Century regulations, the removal of refractory sulfur compounds – particularly dibenzothiophene (DBT) and dimethyl-substituted dibenzothiophenes (DMDBT) – becomes critically important [4,2]. Structural features, namely steric hindrance around the sulfur atom, make hydro-reduction of these molecules prohibitively difficult. A 7-fold increase in catalytic activity [5] or a 5-15 times larger reactor would be required to achieve the stipulated outlet concentration, adding unsustainable cost and energy demands to the process.

A number of technologies have been evaluated as potential supplements to a typical HDS unit, all of which focus on low-

temperature thiohenic aromatic removal. Possibilities include Biodesulfurisation and Oxidative Desulfurisation, both of which require irreversible chemical reaction between sulfurous molecules and the processing medium, large fresh reagent feeds, and extensive process modification for implementation. The ADS mechanism discussed in this study has been shown to enable selective removal of thiophenes while avoiding similar shortcomings[6].

Much of the existing research into ADS focuses on the use of zeolitic structures as candidate adsorbents due to their proven regenerative capabilities and structural flexibility[7]. Titanium-based Nanotubes, synthesised from TiO_2 nanoparticles using the hydrothermal method first presented by Kasuga[8], are a promising adsorbent for industrial desulfurisation. A combination of unique physicochemical properties, high concentration of surface ions, and process scalability[3] mean TiNTs could be a commercially viable adsorbent for use in industrial ADS. Hydrothermal synthesis has also exhibited superior yields compared to more complex procedures[3].

TiNTs present a hollow, open ended, fibre-like morphology similar to that of the carbon nanotubes initially discovered by Ijima[9], with a near-constant internal diameter ranging from 2-10nm[10,11] and a typical aspect ratio greater than 10[3]. During synthesis, between 3 and 8 TiO_2 nanosheets combine laterally and are rolled to produce a tubular morphology with a multi-layered annulus. Interplanar separations of 0.8nm are commonly observed[10,12,11,13]. Moisture content studies concluded that the majority of adsorption occurs on the nanotube surface, while a small degree occurs within the hollow tube. Intercalation rarely contributes significantly to the observed adsorption. Hydrothermal synthesis produces nanotubes with a BET surface area between 200 and 400m^2/g[8,13], which is 33 - 66% of the approximately 600m^2/g possessed by Y zeolites[14]. Lower surfaces areas are attributed to a lack of hierarchical porous structure in TiNTs – a key differentiating factor between nanotubes and previously investigated sorbents.

Expansive research has already been completed on the use of zeolitic structures – particularly those exchanged with transition metal ions – as desulfurisation sorbents. Studies by Kim et al and Tian et al[15,16] have proposed adsorptive capacities and mechanisms for the removal of thiophenes from mixed hydrocarbon feeds. A wide literature search indicated that batch equilibrium capacity can range between 0.02mg/g for exchanged alumni-silicate structures and 22.2mg/g for bimetal exchanged Y zeolites[17,18,19]. In many cases, adsorption is attributed to π-complexation interactions between the sorbent and thiophenic aromatic, although both direct σ S-M bonds and acid-base interactions have been proposed for some adsorbents[20]. π-complexation and σ S-M have been appraised as potential drivers for TiNT behaviour.

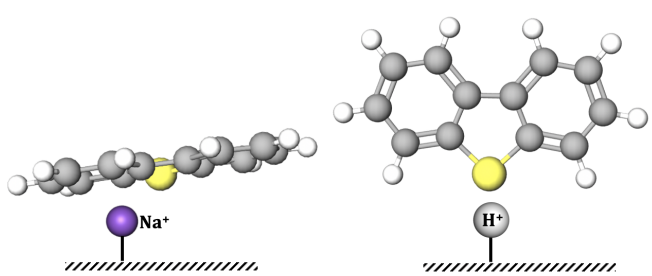


Fig. 1: Possible adsorption mechanisms: π-complexation (left), σ S-M (right)

A 2016 study by Rendon-Rivera[21] provided a preliminary assessment of TiNTs applied to desulfurisation. Room temperature adsorptive capacity under batch conditions was found to be 6.8mg/g for DBT and 2.5mg/g for 4,6-DMDBT, which is 1.7-1.8 times above amorphous TiO_2 and competitive with zeolite capacities found under similar conditions. Despite these promising results, the author presented limited insight into the motivators of effective performance, namely operating conditions and the effect of surface ion exchange. This work builds upon the theory developed for zeolitic materials and presents hypotheses for the predicted behaviour of TiNTs when subject to a range of operating scenarios.

Sun and Li[22] first demonstrated the ability to alter the surface composition of titanate nanotubes using ion exchange. A range of transition metal ions including Cu^{2+} were found to be tightly bounded to the crystal lattice and evenly distributed across the surface of the nanotubes after mixing with an aqueous basic solution. Copper and cerium ion-exchanged zeolites have previously shown to have a considerable impact on the adsorptive capacity and selectivity for sulfur compounds in the presence of aromatics. Notably, a co-exchanged CuCe-Y zeolite was found to have an increased selectivity for 4,6-DMDBT of up to 10 times compared to Ce-Y exchanged zeolites (reference)—. Similar results could be attainable with TiNTs to enhance the adsorptive removal of DBT and 4,6-DMDBT from hydrocarbon fuels. Utilising adsorption via contact with ion-exchanged nanotubes as an initial desulfurisation technology could markedly increase the efficiency of any downstream HDS units. There is therefore considerable industrial motivation to explore the performance of cerium and copper-exchanged TiNTs in the selective removal of DBT and alkyl-DBT molecules.

III. Methods and Materials

A. Reagents and Solvents

All reagents, metal precursors, and organic solvents were purchased from Sigma-Aldrich and used as received. Components used include: Sodium Hydroxide (NaOH, 99.99% purity), Titanium (IV) Oxide, anatase (TiO_2, 99.7% trace metals basis), Copper (II) Nitrate hemi(pentahydrate) ($CuN_2O_6.2.5H_2O$, 98+%), Cerium (IV) Oxide nanopowder (CeO_2, 99.95% trace rare earth metals basis), Zeolyst NaY Zeolite (CBV100 Product, 13.0% Na_2O by wt.) Dibenzothiophene ($C_{12}H_8S$, 98%), 4,6-Dimethyldibenzothiophene ($C_{14}H_{12}S$, 97%), n-Dodecane (99+%), and Hexadecane (anhydrous, 99+%).

B. Model Fuels

The composition of the model diesel fuels used to study the adsorptive and selective performance of TiNTs are detailed in Table I. Quantities of DBT and 4,6-DMDBT - a sterically-hindered refractory compound - mixed with n-dodecane solvent were chosen to reflect the typical sulfur and aromatic content of realistic diesel fuels. Before use, mixtures were

TABLE I: Composition of Model Fuels

	DBT ppmw S	4,6-DMDBT ppmw S
MF1	348	–
MF2	–	302
MF3	348	347

magnetically stirred at 600rpm for 24 h to ensure the aromatic compounds fully dissolved in the n-dodecane.

C. Nanotube Synthesis

Kasuga's alkaline hydrothermal synthesis technique is the most well-documented and reliable method of TiNT production[8]. The typical synthesis route involves the addition of 11g of anatase nanopowder (TiO_2, 99.7%) and 185mL of 10M Sodium Hydroxide (NaOH) to a 200mL PTFE-coated liner and subsequent magnetic stirring for 2 h. The liner is then deposited in a steel autoclave and heated at 120°C for 24 h. Following heating, the resulting slurry was washed with deionised water and filtered to reduce the solution pH to 7. Finally, all remaining product was dried overnight at 120°C to remove residual water. Nanotubes produced in this manner have surfaces populated with Na^+ ions and are labelled Na-TiNT.

In addition to synthesised nanotubes, a number of pre-prepared batches were also investigated due to quantity constraints. Some of the provided batches had been treated with 0.1M H_2SO_4 to exchange Na^+ surface ions with H^+ ions – these samples were denoted as H-TiNT.

D. Batch Adsorption

Batch adsorption experiments were carried out by first adding 1mL weighed volume of model fuel to 10mg of adsorbent in a 1.5mL vial. Mixtures were subsequently sonicated for 5 minutes using a Branson 1510 Ultrasonic Cleaner to increase solid dispersion and assist in aggregate breakdown. Samples were placed inside a Lenton air-circulated oven at constant temperature and subject to rotary agitation at a speed of 40rpm for 2 h to reach equilibrium. Following this, spent adsorbent was separated from the fuel using a 0.45μm StarLab syringe filter to halt adsorption in preparation for processing and analysis in the Gas Chromatographer (GC).

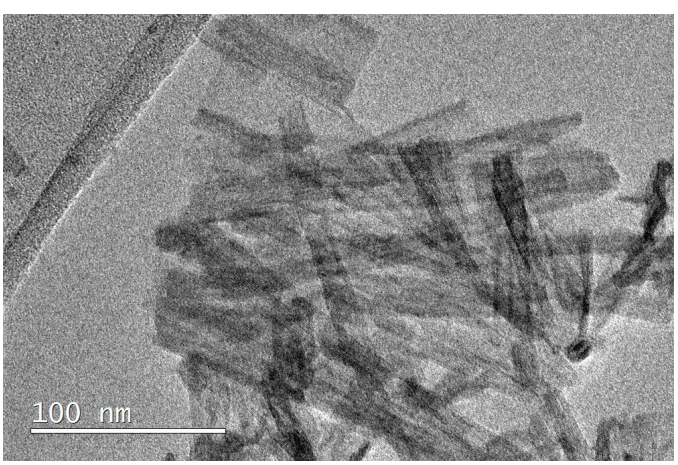


Fig. 2: TEM image of synthesised H-TiNTs

E. Chromatographic Analysis

Prior to analysis, 0.5mL samples of the model fuel were extracted and shaken vigourously with 1mL of n-dodecane (99+%) in a glass vial, which also contained 0.501wt% of hexadecane (99+%) to serve as an internal standard. An internal standard was introduced to reduce measurement errors and introduce uniformity between different experiments. The masses of each compound added were recorded for all samples. Post-adsorption sample processing was completed using a Shimadzu 2014 GC equipped with a 2014-FID flame ionisation detector (FID) and an AOC-20i Auto Injector column.

Fuel-specific calibration curves were first developed to enable accurate conversion between GC spectra and sample composition by mass. All runs were also standardised according to the mass of inputted fuel to enable the calculation of accurate adsorptive capacities. Capacity values were calculated using equations 1 to 4 applied in sequence:

$$m_{adsorbate} = m_{standard} \times \frac{A_{adsorbate}}{CF \times A_{standard}} \tag{1}$$

$$m_{equilibrium} = m_{adsorbate} \times NF \tag{2}$$

$$Adsorption\,\% = \frac{m_{fuel} \times C_{fuel} - m_{equilibrium}}{m_{fuel} \times C_{fuel}} \tag{3}$$

$$q_e = \frac{(m_{fuel} \times C_{fuel}) - m_{equilibrium}}{m_{adsorbent}} \tag{4}$$

A_i refers to the GC peak area of the adsorbate or standard; m_i the mass of either the adsorbate in the post-adsorption sample, the standard, the model fuel, or the solid adsorbent (all mg); C_{fuel} the model fuel concentration of refractory compound (g/g); CF the compound-specific Calibration Factor; and NF the volume Normalisation Factor. q_e represents the adsorptive capacity in $mg/g_{adsorbent}$. Calculated capacities are based upon an averaged GC spectra.

F. Ion Exchange

This study adapted the ion exchange method employed by Viana et al[23] to decorate nanotube surfaces with copper ions. It was previously reported that the H-TiNT form is considered the most promising candidate for ion exchange and was used in this investigation[24]. 200mg of H-TiNT was first added to 200mL of aqueous solution containing 0.05M $CuN_2O_6.2.5H_2O$ before undergoing magnetic stirring for 24 h. Following this, the product was repeatedly washed with deionised water and centrifuged at 3000rpm to remove soluble ions. The resulting slurry was dried at under vacuum at room temperature for 12 h to produce Cu-TiNT powder. A Cerium-exchanged sample, denoted Ce-TiNT, was also developed following an identical method using $(NH_4)_2Ce(NO_3)_6$ precursor.

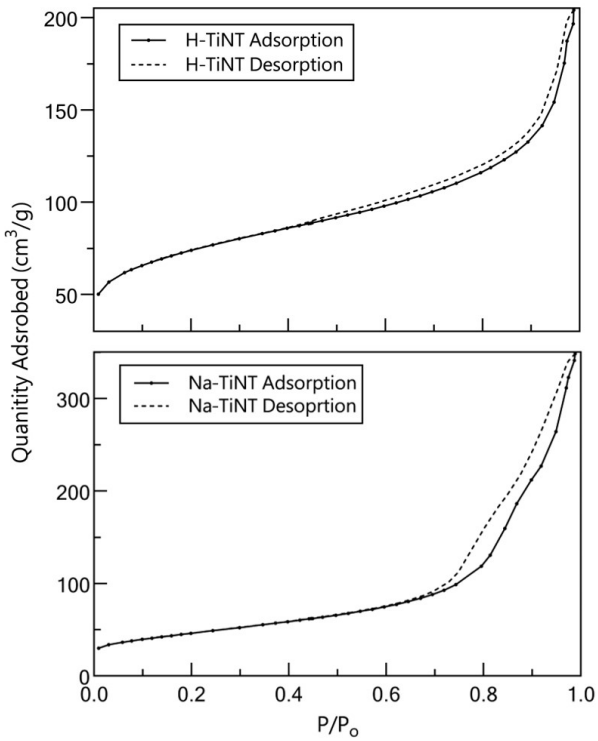


Fig. 3: Nitrogen isotherms: H-TiNT (top), Na-TiNT (bottom)

G. Effect of Drying

To evaluate the effect of drying on the adsorptive performance of TiNT samples, 1.5mL vials containing 10mg of adsorbent were loosely covered with foil and placed inside an air-circulated oven at 120°C. After 18 h, the samples were removed and promptly covered. Following preliminary investigation with undried samples, Batch adsorption experiments were carried out exclusively using the dried nanotubes with MF1 and MF2.[25].

H. Characterisation

Thermogravimetric Analysis (TGA) was performed using a Nezsch STA 449 F5 Jupiter to determine the thermal stability and weight degradation of the synthesised nanotubes. Samples were heated at a rate of 20°C/min from room temperature and subsequently held at a maximum temperature of 600°C for 5 minutes before undergoing cooling to 20°C at a rate of 5°C/min. Nitrogen, used as a purge and protective gas, was supplied at a flow rate of 60ml/min and 20L/min, respectively. The nanotubular structure was fully characterised using a Micromeritics Tristar 3000 with nitrogen adsorption at 77K. Specific surface areas were measured using the Brunauer–Emmett–Teller (BET) method and the total pore volume was evaluated using Barrett-Joyner-Halenda (BJH) analysis. Surface morphology was characterised using Transmission Electron Microscopy performed using a JEOL 2100Plus microscope operated at 200kV. A Thermo K-Alpha device, equipped with an Al Kα source gun, was used to collect X-ray Photoelectronic Spectroscopy (XPS) data and characterise the elemental composition of sample surfaces.

IV. Results and Discussion

A. Characterisations

1) Nitrogen Adsorption: Both Na-TiNT and H-TiNT possess a Type IV nitrogen adsorption isotherm with a H3 hysteresis loop indicating the presence of non-rigid aggregate particles and a mesoporous structure consisting of slit shaped pores[26]. Bavykin and Tsai[11,13] also observed a mesoporous structure for Na-TiNTs and noted that the Nitrogen adsorption isotherms presented features of a H1 hysteresis loop at relative pressures between 0.5 and 0.8. Similar observation were made for the TiNTs used in this study. Figure 3 shows that the H1 hysteresis - defined by the characteristic narrow closing of the hysteresis loop - occurs between relative pressures of 0.4 and 0.7 for H-TiNTs and 0.6 to 0.8 for Na-TiNTs. Such behaviour is indicative of the presence of pores open at both ends, which could correspond to microchannels inside the hollow nanotubes[26].

The BET surface area determined for the synthesised H-TiNTs was 261.4m^2/g, compared to 163.7m^2/g for Na-TiNT and just 50m^2/g for amorphous anatase nanopowder[21]. It can be suggested that a reduction in the sodium content of TiNTs precipitated an increase in surface area. Similar results were reported by Morgado[27], who noted that structural changes caused by the replacement of Na^+ ions with H^+ with acid washing increased the average internal diameter and lowered the skeletal density - both of which contribute to a larger surface area. Despite this, pore distribution data from the BJH desorption shows that Na-TiNTs had a larger total porosity of 0.544cm^2/g compared to 0.259cm^2/g for the H-form, suggesting that the reduced density of H-TiNTs primarily motivated the increased surface area. The Na and H-form nanotubes studied were synthesised in different batches, which could lead to several variations in their final structural properties, as the formation of nanotubes is greatly influenced by synthesis conditions[3]. A hydrothermal treatment temperature close to 130°C leads to nanotubes with a larger porous volume as found by previous work.[11,13]. Tsai et al[13] found that increasing the concentration of acid wash resulted in a 32% decrease in surface area. Total pore volume was also significantly decreased. Considering this previous

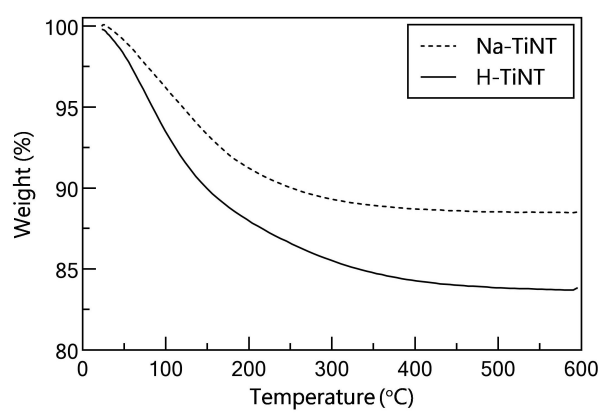
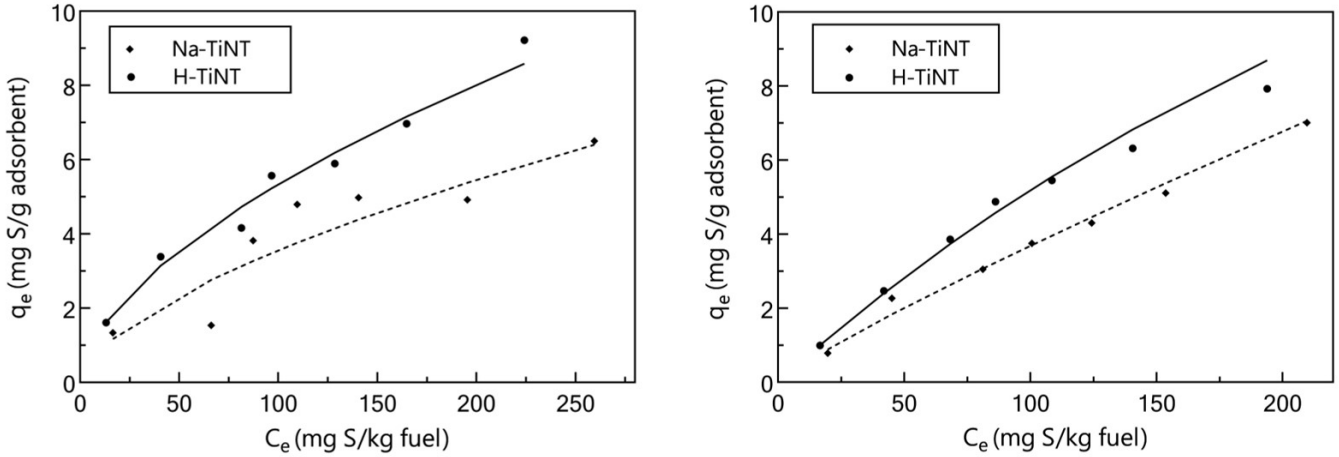


Fig. 4: TGA and Thermal Stability results



(a) Adsorption isotherm using MF1 for Na-TiNT and H-TiNT at 20°C

(b) Adsorption isotherm using MF2 for Na-TiNT and H-TiNT at 20°C

Fig. 5: Experimental isotherms with best-fit model adsorption curves

investigation, it is possible that the larger surface area but lower pore volume of the H-TiNTs compared to Na-TiNTs was caused by the combined effect of a lower hydrothermal synthesis temperature and a high acid wash concentration. An additional possibility is the formation of artificial porous structures caused by agglomeration in the solid sample.

2) Thermal Stability and Weight Degradation Analysis: Thermogravimetric analysis revealed that H-TiNTs have a higher water content compared to Na-TiNTs - as expected from previous studies on the weight degradation of titanate nanotubes [27,22] . The weight-loss between 30°C-120°C was 7.75% and 4.79% for H-TiNTs and Na-TiNTs, respectively. Protonated nanotubes hold a greater amount of adsorbed surface water that is released primarily in this temperature region - as reflected in the comparatively large volatilisation observed. Contrastingly, Na-TiNTs lost a greater proportion of water mass between 120°C - 250°C, which is attributed to the loss of crystallographic water from the nanotubes [25]. The removal of this intercalated water contributed 41% of the total weight-loss observed for Na-form nanotubes, compared to 31% for the H-form. Morgado observed similar results and concluded that increasing the sodium content of nanotubes increased the amount of interlayer water held [27]. Above 400°C, both forms of nanotubes expressed minimal water loss, indicating a loss of thermal stability and a transition of the morphology of samples from nanotubular to a rutile form [25].

3) XPS Analysis: XPS outputs provided the composition and concentration of surface ions for the various forms of TiNT studied. Na-form nanotubes contained 6.07% sodium atoms compared to 2.74% for the H-form. Given the limitations in XPS measurements, it is assumed this arose during the replacement of Na^+ ions with H^+ ions during acid washing. Cerium-exchanged titanate nanotubes contained 1.65% and 1.49% of Na and Ce, respectively. Copper ion exchange performed on Na-TiNT resulted in a surface Cu ion content of 3.29% and a Na ion content of 3.03%. Valance analysis revealed that 70% of all Cu ions present were in the cupric Cu^{2+} form, with the remainder as cuprous Cu^+ ions.

TABLE II: Capacity of titanate nanotubes and Y zeolite

Capacity (mgS/g)	Na-TiNT	H-TiNT	Na-Y
DBT	6.50	9.22	6.34
4,6-DMDBT	7.01	7.93	5.35

B. Adsorptive Performance

1) Equilibrium Capacities: Equilibrium capacities - determined at 20°C - and BET surface areas for Na and H-form nanotubes are presented in Table II alongside comparative values for anatase TiO_2 powder and NaY zeolite. H-TiNT performance is most promising with a capacity of 9.22mgS/g for DBT, while Na-TiNT performance is inferior at 6.50mgS/g for DBT. While differences in surface area could drive such differences, an equally-important factor is likely to be the adsorption strength each surface ion facilitates. H^+ ions are proposed to be more suitable for select surface-adsorbate interactions thus improving capacity, possibly due to their high charge density [20]. Adsorptive performance is 14.1% less for H-TiNTs with 4,6-DMDBT but 7.8% greater for Na-TiNTS with 4,6-DMDBT. A moderate reduction in adsorptive capacity, as highlighted by H-TiNT, is expected. Given the marginally larger projected size of methyl-substituted thiophenes, it is assumed that capacity will be slightly diminished for 4,6-DMDBT molecules. It is possible that experimental error or non-uniform adsorption has resulted in an increase in the Na-TiNT sample. H-TiNTs presented a more radical reduction than predicted, which could imply H-form nanotubes rely on a different adsorption mechanism that is explicitly hindered by alkyl substitution. Section IV-C presents two possible mechanisms for adsorption, although wider investigation is needed to understand the behaviour of H surface ions.

Both TiNT samples showed competitive performance relative to the tested anatase and Y zeolite products. Surface area differences are likely drivers of the 70-80% DBT capacity

increase gained from the titanate structural change[21]. The inferior performance exhibited by the Y zeolite is likely due to the selected adsorption temperature, which is considerably below normal zeolite operating conditions[28,29]. Discussion of the effects of temperature on mass transfer is presented in Section IV-E. Comparison of obtained capacities to compiled equilibrium capacities presented by Lee[20] confirmed that TiNT adsorbents possess a competitive adsorbent capacity and room temperature, which is a notable advantage over many zeolitic adsorbents.

Adsorptive capacities calculated on a surface area basis reinforce this finding. For DBT, samples presented capacities of $1.10 \mu mol/m^2$ and $1.24 \mu mol/m^2$ for H-TiNTs and Na-TiNTs, respectively. Once again, this defies predicted behaviour. It is likely that variations in the synthesis process of different batches has caused this, as small changes markedly impacts surface areas[3].

2) Isotherm Modelling: Following linear fitting of several adsorption models to experimental data, the best fit was determined through comparison to experimental data and calculation of R^2 correlation coefficients. Freundlich isotherm models provided the most appropriate fit for all adsorption processes apart from the adsorption of 4,6-DMDBT on H-TiNT, which fitted best with the Langmuir isotherm. Table III presents the obtained isotherm constants and R^2 values, while Figure 5 shows the experimental data plotted alongside the fitted theoretical model with the largest R^2 value . The value of n for each fitted Freundlich isotherm is greater than 1, which is indicative of favourable adsorption of DBT and 4,6-DMDBT on TiNTs[30]. Significantly, the greater K_f values for H-TiNT implies superior adsorptive desulfurisation capacity compared to the Na-form.

The heterogeneity in surface energies assumed by the Freundlich isotherm model is expected given the presence of multiple different surface ions highlighted in Section IV-A3. The model fuels used to create each isotherm, MF1 and MF2, only contain one type of thiophenic compound, meaning heterogeneity must arise from the presence of adsorption sites with varying energies. The Na^+ and H^+ ions present on the surface of both the Na-form and H-form nanotubes are therefore expected to contribute heterogeneous adsorption sites.

Developed theory covering the adsorptive mechanisms used in thiophene adsorption is based upon monolayer adsorption. Consequently, the multilayer assumption of a Freundlich model should not be applicable to this adsorption scenario. However, knowledge presented by Ayawei et al shows that a Freundlich isotherm is most suitable for low adsorbate concentrations and intermediate surface coverages[31]. To confirm tested samples lie within this categorisation of surface coverage, approximated coverage values were calculated using an estimated projected area for DBT. Following Ramirez-Cabrera et al, a carbon atom was assumed to cover $0.026 nm^2$ of an adosrbent surface[32]. Calculations show that, for each m^2 of area, 25.3% is occupied by DBT molecules in H-TiNT samples and 28.5% is occupied in Na-TiNT samples. These values lie clearly within the intermediate coverage range, verifying the determined Freundlich models. Further experimentation with sulfur concentrations closer to the saturation capacity of the nanotubes is needed to validate isotherm models. It is predicted that these tests will result in the expected characteristic plateau of the Langmuir isotherm being observed.

TABLE III: Isotherm parameters

Model	Parameters	Na-TiNT		H-TiNT	
		MF1	MF2	MF1	MF2
Freundlich	K_f (mg/g)	0.251	0.075	0.424	0.114
	n	1.62	1.14	1.70	1.18
	R^2	0.920	0.995	0.987	0.990
Langmuir	K_L (kg/mg)	0.025	0.000867	0.0227	0.00233
	q_m (mg/g)	5.38	56.4	8.77	31.2
	R^2	0.884	0.994	0.936	0.995
Temkin	K_t (kg/mg)	0.117	0.0675	0.147	0.0810
	B_t	1.88	2.39	2.43	2.79
	R^2	0.892	0.955	0.940	0.979

C. Proposed Mechanism

Mechanistic analysis of adsorption onto zeolitic sorbents has been considered in detail. It is suggested that theory developed for these molecular interactions can be extended to TiNT adsorption.

There is no reported evidence of chemisorption occurring between sulfurous aromatics and adsorbent surfaces, likely due to the steric hindrance and charge distribution present in many thiophenes. Instead, previous studies have widely reported enhanced physisorptive effects as the primary sorption mechanism. Bond energy calculations[33] and FTIR spectroscopy[34] have both suggested that a supplementary adsorption mode drives interaction. π-complexation is proposed as the dominating mechanism for aromatic adsorption on surfaces. Molecular Orbital calculations by Takahashi et al[35] have concluded surface metal ions can engage in charge transfer with aromatic rings and produce relatively strong physisorptive effects. This study hypothesises that similar interactions drive DBT and 4,6-DMDBT adsorption on TiNT surfaces. As all aromatic rings are affected by the same bonding mechanism, competitive adsorption is a notable limiting feature of operation.

Additionally, evidence suggests that select surface ions can form direct σ S-M bonds with sulfur atoms[20]. The prominence of such interactions depends on the charge number and size of surface cations, steric hindrance around the thiophenic sulfur atom, and surface loading factors. Both H^+ and Lanthanide ions are known to preferentially utilise this mechanism. Depending on their strength, S-M bonds can significantly improve selectivity towards sulfur-containing molecules. Previous investigations have already shown that certain surface ions can result in distinct increases in adsorption of sulfur compounds despite competition from aromatics present[36,16]. Furthermore, should an adsorbent possess surface ions that are better suited to S-M interaction - as is the case H-TiNT samples tested - the steric hindrance of thiophenes can have an inflated effect on capacity. With limited π-complexation capabilities, some adsorbents rely more exclusively on S-M interactions that can easily be limited by

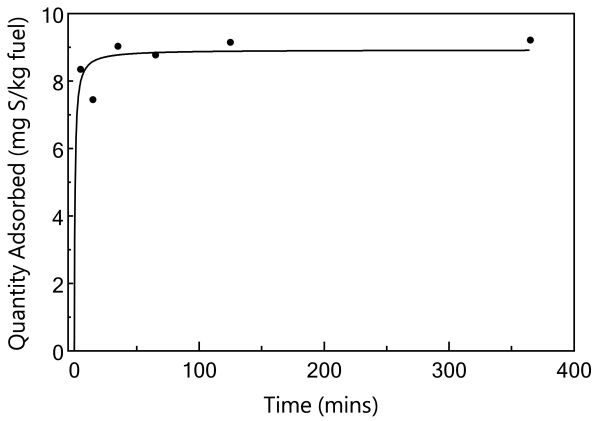


Fig. 6: Equilibrium data for H-TiNT

steric factors. It is theorised that such limitations resulted in the noticeable drop in H-TiNT capacity between DBT and 4,6-DMDBT adsorption found in this study.

Surface ion exchange plays a major role in altering the observed mechanisms and consequently impacting both sorbent performance and selectivity. A preliminary assessment of these effects is presented in Section IV-H.

D. Equilibrium

Evaluation of the time taken to reach equilibrium showed an overwhelming proportion of the adsorption process was completed within 5 minutes of adsorbate addition. Figure 6 highlights that 90.6% of adsorption on H-TiNTs was completed in this period and desulfurisation occurring in the remaining time was relatively insignificant. Both findings validate the choice of a 2 h adsorption time and provide insights into the effect of diffusional limitations on the process.

The fast adsorption indicated by Figure 6 suggests the impact of mass transfer resistances induced by nanoporous structures are limited. This is confirmed by TEM imagery in both this study and others[37,38], which show the microporous structure within TiNT particles is more limited than the hierarchical structure observed in zeolites[39]. Desulfurisation processes utilising Y-zeolites have found that 90% of equilibrium capacity was obtained following 240 minutes – a notable increase on the results observed[40]. It is proposed that adsorbate molecules can quickly access available sites and interact with surface ions, facilitating rapid initial rates[41]. Following surface occupation, it is postulated that the sub-nanometre interlayer spaces[42] and the resulting steric hindrance of large aromatic compounds – as well as low adsorption temperatures – render the rate of diffusion to interlayer surface sites almost negligible[43]. Consequently, the residual increases in capacity observed beyond the first 5 minutes are attributed to the gradual occupation of remaining surface sites, as opposed to adsorption onto sites within the nanotube porous structure.

E. Temperature Effects

Operating temperature plays a key role in determining the adsorptive capacity of many sorbents, particularly those that rely primarily on physisorptive interactions. Typically, heightened diffusion rates and enhanced equilibrium capacities drive increases in adsorptive capacity as process temperatures are raised above ambient conditions[44]. Contrary to conventional observations, experiments carried out at 20°C, 30°C, 40°C, 50°C, and 60°C for H-TiNT adsorbents produced a declining equilibrium capacity beyond 20°C. Adsorptive capacity is reduced by 6.9% for H-TiNTs, indicating that the traditional benefits gained from elevated temperatures do not apply to nanotube adsorption. Such a trend is observed in contrast with zeolite adsorbent behaviour, which Velu et al found to express superior performance at or above 80°C[28,29].

It is theorised that TiNT adsorbents can operate effectively at ambient temperatures due to a limited hierarchical pore structure and consequently lower diffusional barriers relative to other structures. With adsorption expected to occur exclusively on the outer particle surface, the adsorbate faces minimal intraparticle resistance to mass transfer and is replaced as the rate-limiting step by external film diffusion[45]. As temperature increases, the lack of mass transfer resistance raises the likelihood of molecule desorption – a phenomenon that is emphasised by the absence of chemical bonds with the surface. Therefore, while adsorption rate is elevated at higher temperatures, desorption rates are also accelerated to produce a moderately diminished adsorptive capacity. To evaluate the ease of desorption under the circumstances considered, adsorption energy should be determined through Microcalorimetry.

F. The Effect of Drying

Figure 8 shows that drying nanotubes prior to adsorption resulted in an appreciable improvement to the adsorptive capacity for nanotubes for both DBT and 4,6-DMDBT. Drying had the greatest effect on H-TiNTs - resulting in a 32.4% higher adsorptive capacity for DBT and 30.7% for 4,6-DMDBT. More water was removed from the H-TiNTs from low temperature drying at 120°C due to the larger amount of adsorbed surface water present on its surface - a theory verified by TGA readings presented in Section IV-A2. The removal of this water allowed for more sulfur compounds to

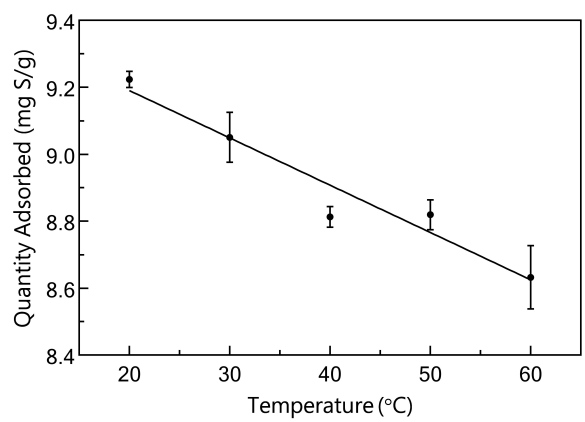


Fig. 7: Adsorption capacity of H-TiNT as a function of temperature

be adsorbed onto the surface of H-TiNTs compared to Na-TiNTs. Na-form nanotubes experience less pronounced effects because a significant proportion of adsorbed water is stored as intercalated water. Increasing the drying temperature to 250°C would release crystallographic water. Despite this, it is postulated that the adsorption of sulfur compounds occurs primarily on nanotube surfaces and not in the interlayer gaps. Therefore, the 16.3% increase in adsorptive capacity achieved for 4,6-DMDBT is assumed to be a maxmium for the Na-TiNT samples. Surprisingly, the performance of the dried sodium titanate nanotubes was 3.1%worse than the orignal samples for the removal of DBT from MF1. There is no evidence to suggest why this trend would be observed and the anomalous result is likely due to spurious experimental variations.

G. Selectivity

Selectivity towards adsorption of refractory compounds is a core performance metric for potential desulfurisation adsorbents. As mentioned in Section IV-C, previous work has shown the prohibitive impact including non-sulfurous aromatics has on some adsorption processes. In this case, selectivity between two refractory compounds has been considered and the results presented in Table IV. It was observed that, at equilibrium, 8.66mg/g of DBT and 8.46mg/g of 4,6-DMDBT were adsorbed onto the surface of H-TiNTs – equivalent to $1.03\mu mol/m^2$ and $1.01\mu mol/m^2$, respectively. These results imply a moderate preference towards DBT adsorption, with a 51%-49% split between DBT and 4,6-DMDBT attachment. The similarities in adsorption mechanism for both thiophenic compounds motivate the conclusion that this is a predictable result[46,47].

Theory initially presented by Jiang et al[40] proposed a decrease in adsorption energy with molecule size, although it was subsequently concluded that the microporous zeolitic structure of the sorbent used drove behaviour seen in this study[20]. On the other hand, it has been suggested that substituted aromatic rings express enhanced π-complexation, facilitated through electron density donation from substituted groups to the Benzene ring[48]. Comparable behaviour was not observed in this study, which could be an impact of steric hindrance limiting 4,6-DMDBT surface affinity[15].

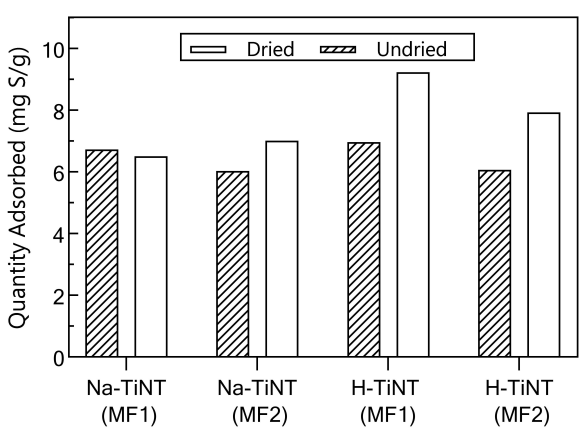


Fig. 8: Effect of drying on Na-TiNT and H-TiNT

TABLE IV: Selectivity for mixed fuel MF3

	DBT mgS/g	4,6-DMDBT mgS/g	Selectivity Ratio
H-TiNT	8.66	8.46	51:49
Na-TiNT	7.59	7.60	50:50

H. Ion Exchange

Ion exchange processes are generally considered a promising method of improving both selectivity towards thiophenes in mixed feeds and increasing the equilibrium capacity of DBT and 4,6-DMDBT for a range of adsorbent materials[46,47,20]. In an extension to existing research, which focused on ion exchanged Y zeolites, both Cu^{2+} and Ce^{4+} ions were exchanged onto H-TiNT samples using the method described in Section III-F. Table V presents the equilibrium capacities achieved by four samples of nanotubes, each supporting a different surface ion. Copper ion exchange produced the highest capacity, although this represented just a 0.52% increase compared to H-TiNT values. The substitution of Cu^{2+} ions, which possess similar Lewis Acid properties to H^+ ions, appears to strengthen the π-complexation interaction between the surface ion and aromatic rings in the adsorbate. Strengthening is achieved through back-donation from occupied d-orbitals of the copper cations – such behaviour cannot be expressed when using sodium or hydrogen-form nanotubes[35]. It was also noted that charge transfer mechanisms with Cu^{2+} surface ions improve selectivity towards refractory compounds, while similar behaviour is not attainable without transition metal-exchanged nanotubes.

XPS readings presented in Section IV-A3 offer a possible explanation for the comparatively insignificant capacity change noted above. 3.29% of the Cu-TiNT sample surface consists of Cu ions - the majority of which are cupric Cu^{2+}. At noted by Hernandez-Maldonado et al[46], π-complexation os only facilitated by cuprous Cu^+ ions. Such a limited proportion of Cu^+ on the Cu-TiNT surface indicates the adsorptive capacity is not maximised. Changes to the treatment and synthesis procedures can produce a greater proportion of reduced ions in future samples.

Cerium-exchanged nanotubes exhibited a moderate increase in performance compared to Na-TiNTs, with an equilibrium capacity of 7.85mgS/g. While this figure is 20.7% above Na-TiNT performance, limitations in the synthesis technique used imply this figure could be increased further using a modified approach. XPS results emphasise these concerns: an exchanged ion concentration of 1.49% is appreciably lower than the value obtained in previous studies[23]. Highly-positive Lanthanide ions such as Ce^{4+} possess the unique ability to form direct σ bonds with sulfur atoms while adsorbed onto a nanostructure - denoted as σ S-M[28]. Such interaction provides significant selectivity towards sulfurous molecules, although the confinement of the sulfur atom in DBT and 4,6-DMDBT reduces the bonding effectiveness compared to π-complexation. Hence, Ce-TiNTs demonstrated a 15.3% lower capacity than Cu-TiNTs. It was noted that performance of pure Cerium Nanowires (CeNW) – a product derived from Cerium

TABLE V: Impact of Surface Ion on DBT Capacity

	H-TiNT	Na-TiNT	Cu-TiNT	Ce-TiNT	CeNW
DBT mgS/g	9.22	6.50	9.27	7.85	6.52

(IV) Oxide nanoparticles – was 6.52mgS/g, appreciably lower than for surface exchanged Ce-TiNTs. This behaviour is likely due to the importance of surface ions compared to the underlying material and the morphology of nanowires, which may be unsuitable for the adsorption of large aromatics.

While extensive investigation into the presented methods are available for Y zeolites, the current scope for TiNTs is more restricted. Consequently, it is recommended that supplementary practical tests and Density Function Theory models are completed to fully understand the nuances of TiNT mechanisms. Previous works [18,28,14] have also highlighted the potential of bimetal-exchanged adsorbents, which express synergistic behaviour that enhances both capacity and selectivity.

V. Future Work

The findings of this exploratory investigation provide strong motivations for further targeted research into the desulfurisation performance of TiNTs. Crucially, a detailed study of the adsorption mechanisms using techniques such as Density Function Theory would determine the validity of the proposed π-complexation and σ S-M adsorption mechanisms. In order to also verify surface behaviour, higher concentrations must be tested to fully develop the proposed adsorption isotherms.

To verify industrial applicability, continuous adsorption tests should be performed to determine the full influence of operating conditions on saturation capacities of the adsorbents. Although not mentioned in detail above, investigations into the desorption and regeneration properties of TiNTs are essential considerations. Finally, insights into the efficient pelletisation of nanotubular structures is a vital element of process development. Advances in each of these areas is needed to fully evaluate TiNT adsorption as a breakthrough SARS technology and efficient supplementary process to HDS.

VI. Conclusion

New insights into the performance nuances of titanate nanotubes when applied to Adsorptive Desulfurisation processes have been successfully developed in this study. By strategically extending the accepted understanding of zeolitic adsorbent performance to nanotube behaviour, numerous performance drivers could be investigated, including operating temperature, drying effects, equilibrium, and surface ion composition.

Hydrogen-form nanotubes were found to be the most competitive performer, achieving an equilibrium adsorptive capacity of 9.22mgS/g for DBT - equivalent to 0.035mg/m^2 - under ambient operating conditions. At 20°C, both NaY zeolite and amorphous TiO_2 powder produced substantially lower equilibrium capacities - highlighting a key operating advantage of nanotubes. There is substantial evidence to suggest thiophenic compounds adsorb onto nanotube surfaces using the same π-complexation and σ S-M mechanisms observed in zeolite adsorption.

Selectivity investigations confirmed that, while all thiophenes likely share a common adsorption mechanism, steric hindrance and the surface ion can have a moderate effect on selectivity. Ion exchange tests confirmed that cerium and copper addition can be used to improve the selectivity and adsorptive capacity relative to as-synthesised nanotubes. XPS spectra confirmed that limited exchange occurred; future tests at higher surface loadings are needed to fully enumerate the benefits of this process. Furthermore, the possibility of synergistic effects arising from bimetal copper-cerium exchange should be considered in significant detail.

The completed experimentation has yielded an expanded knowledge base in the area of selective sulfur adsorption and validated titanate nanotubes as a candidate adsorbent for sustainable industrial application. Proven scalability and low intensity operating conditions are both key factors boosting the appeal of nanotubes as selective adsorbents. Nevertheless,realistic application would almost certainly require a combination of processes and adsorbents to efficiently remove the myriad of sulfurous compounds present in hydrocarbon fuels.

VII. Acknowledgements

The authors wish to offer their gratitude to Miss Rui Zhang and Miss Yiming Zhang for their support in all areas of this investigation, particularly in adsorbent preparation and sample characterisation. Motaz Khawaji is thanked for providing additional nanotube samples for analysis. Dr Roberto Rinaldi is also acknowledged for his contribution to the sample analysis process employed.

References

[1] E. I. I. U.S., "U.S. Sulfur Content (Weighted Average) of Crude Oil Input to Refineries," 2019. [Online]. Available: https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET{&}s=MCRS1US2{&}f=M

[2] C. Courville, "ULTRA LOW SULFUR DIESEL (ULSD): THE GOOD, THE BAD, AND THE RUSTY," 2018. [Online]. Available: https://www.axi-international.com/ultra-low-sulfur-diesel-ulsd/

[3] D. V. Bavykin, J. M. Friedrich, and F. C. Walsh, "Protonated titanates and TiO2 nanostructured materials: Synthesis, properties, and applications," *Advanced Materials*, vol. 18, no. 21, pp. 2807–2824, nov 2006.

[4] C. Song, "An overview of new approaches to deep desulfurization for ultra-clean gasoline, diesel fuel and jet fuel," *Catalysis Today*, vol. 86, no. 1-4, pp. 211–263, 2003.

[5] T. Chiranjeevi, G. M. Kumaran, and G. M. Dhar, "Synthesis, characterization, and evaluation of mesoporous MCM-41-supported molybdenum hydrotreating catalysts," *Petroleum Science and Technology*, vol. 26, no. 6, pp. 690–703, 2008.

[6] V. Chandra Srivastava, "An evaluation of desulfurization technologies for sulfur removal from liquid fuels," *RSC Advances*, vol. 2, no. 3, pp. 759–783, 2012.

[7] R. Dehghan and M. Anbia, "Zeolites for adsorptive desulfurization from fuels: A review," *Fuel Processing Technology*, vol. 167, pp. 99–116, dec 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378382017301637bb0030

[8] T. Kasuga, M. Hiramatsu, A. Hoson, T. Sekino, and K. Niihara, "Formation of Titanium Oxide Nanotube," *Langmuir*, vol. 14, no. 12, pp. 3160–3163, jun 1998. [Online]. Available: https://doi.org/10.1021/la9713816

[9] Sumlo Iijima, "Helical microtubules of graphitic carbon," *Nature*, vol. 354, no. 354, pp. 56–58, 1991.

[10] E. Lorençon, D. C. Alves, K. Krambrock, E. S. Ávila, R. R. Resende, A. S. Ferlauto, and R. M. Lago, "Oxidative desulfurization of dibenzothiophene over titanate nanotubes," *Fuel*, vol. 132, pp. 53–61, 2014.

[11] D. V. Bavykin, V. N. Parmon, A. A. Lapkin, and F. C. Walsh, "The effect of hydrothermal conditions on the mesoporous structure of TiO2 nanotubes," *J. Mater. Chem.*, vol. 14, no. 22, pp. 3370–3377, 2004. [Online]. Available: http://dx.doi.org/10.1039/B406378C

[12] T. Kasuga, "Formation of titanium oxide nanotubes using chemical treatments and their characteristic properties," in *Thin Solid Films*, 2006.

[13] C.-C. Tsai and H. Teng, "Regulation of the Physical Characteristics of Titania Nanotube Aggregates Synthesized from Hydrothermal Treatment," *Chemistry of Materials*, vol. 16, no. 22, pp. 4352–4358, nov 2004. [Online]. Available: https://doi.org/10.1021/cm049643u

[14] H. Song, H. Song, X. Wan, M. Dai, J. Zhang, and F. Li, "Deep desulfurization of model gasoline by selective adsorption over Cu-Ce bimetal ion-exchanged y zeolite," *Fuel Processing Technology*, vol. 116, pp. 52–62, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.fuproc.2013.04.017

[15] J. H. Kim, X. Ma, A. Zhou, and C. Song, "Ultra-deep desulfurization and denitrogenation of diesel fuel by selective adsorption over three different adsorbents: A study on adsorptive selectivity and mechanism," *Catalysis Today*, vol. 111, no. 1-2, pp. 74–83, 2006.

[16] F. Tian, W. Wu, Z. Jiang, C. Liang, Y. Yang, P. Ying, X. Sun, T. Cai, and C. Li, "The study of thiophene adsorption onto La(III)-exchanged zeolite NaY by FT-IR spectroscopy," *Journal of Colloid and Interface Science*, vol. 301, no. 2, pp. 395–401, 2006.

[17] H. Y. Sun, L. P. Sun, F. Li, and L. Zhang, "Adsorption of benzothiophene from fuels on modified NaY zeolites," *Fuel Processing Technology*, vol. 134, pp. 284–289, 2015. [Online]. Available: https://doi.org/10.1016/j.fuproc.2015.02.010

[18] J. Wang, F. Xu, W. jie Xie, Z. jian Mei, Q. zhuo Zhang, J. Cai, and W. min Cai, "The enhanced adsorption of dibenzothiophene onto cerium/nickel-exchanged zeolite Y," *Journal of Hazardous Materials*, vol. 163, no. 2-3, pp. 538–543, 2009.

[19] X. Xu, S. Zhang, P. Li, and Y. Shen, "Equilibrium and kinetics of Jet-A fuel desulfurization by selective adsorption at room temperatures," *Fuel*, vol. 111, pp. 172–179, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.fuel.2013.04.068

[20] K. X. Lee and J. A. Valla, "Adsorptive desulfurization of liquid hydrocarbons using zeolite-based sorbents: A comprehensive review," *Reaction Chemistry and Engineering*, vol. 4, no. 8, pp. 1357–1386, 2019.

[21] A. Rendon-Rivera, M. A. Cortes-Jacome, E. Lopez-Salinas, M. L. Mosqueira, and J. A. Toledo-Antonio, "Adsorption of nitrogen and sulphur organic-compounds on titania nanotubes," *International Journal of Engineering Research & Science*, vol. 2, pp. 156–166, 2016. [Online]. Available: https://www.academia.edu/28254810/Adsorption{_}of{_}Nitrogen{_}and{_}Sulphur{_}Organic-Compounds{_}on{_}Titania{_}Nanotubes

[22] X. Sun and Y. Li, "Synthesis and characterization of ion-exchangeable titanate nanotubes," *Chemistry - A European Journal*, vol. 9, no. 10, pp. 2229–2238, may 2003.

[23] B. C. Viana, O. P. Ferreira, A. G. Souza Filho, C. M. Rodrigues, S. G. Moraes, J. M. Filho, and O. L. Alves, "Decorating titanate nanotubes with CeO2 nanoparticles," *Journal of Physical Chemistry C*, vol. 113, no. 47, pp. 20 234–20 239, 2009.

[24] T. Gao, Q. Wu, H. Fjellvåg, and P. Norby, "Topological properties of titanate nanotubes," *Journal of Physical Chemistry C*, vol. 112, no. 23, pp. 8548–8552, 2008.

[25] D. V. Bavykin, J. M. Friedrich, and F. C. Walsh, "Chapter 4 chemical properties, transformation and functionalization of elongated titanium oxide nanostructures," in *Titanate and Titania Nanotubes: Synthesis*. The Royal Society of Chemistry, 2009, pp. 85–108. [Online]. Available: http://dx.doi.org/10.1039/9781849730778-00085

[26] K. S. Sing, D. H. Everett, R. A. Haul, L. Moscou, R. A. Pierotti, J. Rouquerol, and T. Siemieniewska, "Reporting Physisorption Data for Gas/Solid Systems with Special Reference to the Determination of Surface Area and Porosity," *Pure and Applied Chemistry*, vol. 57, no. 4, pp. 603–619, 1985.

[27] E. Morgado, M. A. de Abreu, O. R. Pravia, B. A. Marinkovic, P. M. Jardim, F. C. Rizzo, and A. S. Araújo, "A study on the structure and thermal stability of titanate nanotubes as a function of sodium content," *Solid State Sciences*, vol. 8, no. 8, pp. 888–900, aug 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1293255806000896{#}tbl001

[28] S. Velu, X. Ma, and C. Song, "Zeolite-based adsorbents for desulfurization of jet fuel by selective adsorption," *ACS Division of Fuel Chemistry, Preprints*, vol. 47, no. 2, pp. 447–448, 2002.

[29] S. Velu, C. Song, M. H. Engelhard, and Y. H. Chin, "Adsorptive removal of organic sulfur compounds from jet fuel over K-exchanged NiY zeolites prepared by impregnation and ion exchange," *Industrial and Engineering Chemistry Research*, vol. 44, no. 15, pp. 5740–5749, 2005.

[30] A. O. Dada, A. P. Olalekan, A. M. Olatunya, and Dada, "Langmuir, Freundlich, Temkin and Dubinin-Radushkevich Isotherms Studies of Equilibrium Sorption of Zn 2+ Unto Phosphoric Acid Modified Rice Husk," Tech. Rep. 1. [Online]. Available: www.iosrjournals.org

[31] N. Ayawei, A. N. Ebelegi, and D. Wankasi, "Modelling and Interpretation of Adsorption Isotherms," *Journal of Chemistry*, vol. 2017, 2017.

[32] E. Ramírez-Cabrera, A. Atkinson, and D. Chadwick, "Reactivity of ceria, Gd- and Nb-doped ceria to methane," *Applied Catalysis B: Environmental*, vol. 36, no. 3, pp. 193–206, 2002.

[33] K. X. Lee and J. A. Valla, "Investigation of metal-exchanged mesoporous Y zeolites for the adsorptive desulfurization of liquid fuels," *Applied Catalysis B: Environmental*, vol. 201, pp. 359–369, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.apcatb.2016.08.018

[34] Y. Shi, X. Yang, F. Tian, C. Jia, and Y. Chen, "Effects of toluene on thiophene adsorption over NaY and Ce(IV)Y zeolites," *Journal of Natural Gas Chemistry*, vol. 21, no. 4, pp. 421–425, 2012. [Online]. Available: http://dx.doi.org/10.1016/S1003-9953(11)60385-X

[35] A. Takahashi, F. H. Yang, and R. T. Yang, "New sorbents for desulfurization by π-complexation: Thiophene/benzene adsorption," *Industrial and Engineering Chemistry Research*, vol. 41, no. 10, pp. 2487–2496, 2002.

[36] M. L. Oliveira, A. A. Miranda, C. M. Barbosa, C. L. Cavalcante, D. C. Azevedo, and E. Rodriguez-Castellon, "Adsorption of thiophene and toluene on NaY zeolites exchanged with Ag(I), Ni(II) and Zn(II)," *Fuel*, vol. 88, no. 10, pp. 1885–1892, 2009. [Online]. Available: http://dx.doi.org/10.1016/j.fuel.2009.04.011

[37] I. Song, H. Lee, S. W. Jeon, D. ha Lim, and D. H. Kim, "Hydrothermal Synthesis of Titanate Nanotubes with Different Pore Structure and its Effect on the Catalytic Performance of V 2 O 5 -WO 3 /Titanate Nanotube Catalysts for NH 3 -SCR," *Topics in Catalysis*, vol. 62, no. 1-4, pp. 214–218, 2019. [Online]. Available: http://dx.doi.org/10.1007/s11244-018-1085-0

[38] S. X. Lu, H. Zhong, D. M. Mo, Z. Hu, H. L. Zhou, and Y. Yao, "A H-titanate nanotube with superior oxidative desulfurization selectivity," *Green Chemistry*, vol. 19, no. 5, pp. 1371–1377, 2017. [Online]. Available: http://xlink.rsc.org/?DOI=C6GC03573F

[39] T. Yokoi, "Characterization of zeolites by advanced SEM / STEM techniques," *Scientific Instrument News*, vol. 7, pp. 17–23, 2016. [Online]. Available: https://www.hitachi-hightech.com/file/global/pdf/sinews/si{_}report/07017.pdf

[40] J. Jiang and F. T. Ng, "Production of low sulfur diesel fuel via adsorption: An equilibrium and kinetic study on the adsorption of dibenzothiophene onto NaY zeolite," *Adsorption*, vol. 16, no. 6, pp. 549–558, 2010.

[41] H. Kochkar, A. Turki, L. Bergaoui, G. Berhault, and A. Ghorbel, "Study of Pd(II) adsorption over titanate nanotubes of different diameters," *Journal of Colloid and Interface Science*, vol. 331, no. 1, pp. 27–31, 2009. [Online]. Available: http://dx.doi.org/10.1016/j.jcis.2008.11.002

[42] R. Ma, T. Sasaki, and Y. Bando, "Alkali metal cation intercalation properties of titanate nanotubes," *Chemical Communications*, no. 7, pp. 948–950, 2005.

[43] L. Xiong, C. Chen, Q. Chen, and J. Ni, "Adsorption of Pb(II) and Cd(II) from aqueous solutions using titanate nanotubes prepared via hydrothermal method," *Journal of Hazardous Materials*, vol. 189, no. 3, pp. 741–748, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.jhazmat.2011.03.006

[44] L. Nouri, I. Ghodbane, O. Hamdaoui, and M. Chiha, "Batch sorption dynamics and equilibrium for the removal of cadmium ions from aqueous phase using wheat bran," *Journal of Hazardous Materials*, vol. 149, no. 1, pp. 115–125, 2007.

[45] L. Xiong, Y. Yang, J. Mai, W. Sun, C. Zhang, D. Wei, Q. Chen, and J. Ni, "Adsorption behavior of methylene blue onto titanate nanotubes," *Chemical Engineering Journal*, vol. 156, no. 2, pp. 313–320, 2010.

[46] A. J. Hernández-Maldonado, F. H. Yang, G. Qi, and R. T. Yang, "Desulfurization of transportation fuels by π-complexation sorbents: Cu(I)-, Ni(II)-, and Zn(II)-zeolites," *Applied Catalysis B: Environmental*, vol. 56, no. 1-2 SPEC. ISS., pp. 111–126, 2005.

[47] V. M. Bhandari, C. H. Ko, J. G. Park, S. S. Han, S. H. Cho, and J. N. Kim, "Desulfurization of diesel using ion-exchanged zeolites," *Chemical Engineering Science*, vol. 61, no. 8, pp. 2599–2608, 2006.

[48] F. H. Yang, A. J. Hernandez-Maldonado, and R. T. Yang, "Selective adsorption of organosulfur compounds from transportation fuels by π-complexation," *Separation Science and Technology*, vol. 39, no. 8, pp. 1717–1732, 2004.

Global Optimisation and Machine Learning Techniques for Characterising Design Spaces in Pharmaceutical Processes

Hamish Mitchell and James Morrissey
Department of Chemical Engineering, Imperial College London, UK

Abstract

The need for a more structured approach to pharmaceutical process development has been identified in recent years. The concept of a design space is a key aspect of the pharmaceutical Quality by Design (QbD) initiative introduced in the International Conference on Harmonisation (ICH) Q8 guideline for pharmaceutical development. This paper lays out a methodology to effectively characterise design spaces, which can then be used directly in pharmaceutical manufacturing. Mechanistic models are used to generate probabilistic design spaces via Monte Carlo simulation, followed by the use of artificial neural network (ANN) regression to provide an algebraic relationship between critical process parameters (CPPs) and the probability of reaching critical quality attribute (CQA) targets. A relaxation of a semi-infinite global optimisation problem is then used to find upper and lower bounds on each of the CPPs to characterise hyperrectangles which represent the largest operating region within the design space. These operating spaces are advantageous compared to the original design space as they can be communicated to process operators as linear, independent ranges in the CPPs. Two successful examples of the methodology are presented in this paper - a simple batch reactor with two CPPs, and a more complex Suzuki coupling problem with four CPPs.

Keywords: Quality by Design, Design Space, Neural Networks, Optimisation, Pharmaceuticals, Monte Carlo Simulations

1 Introduction

Since the publication of the International Conference on Harmonisation (ICH) Q8 document, quality by design (QbD) - developing processes which inherently produce safe, effective pharmaceutical products - has been of great importance. A crucial aspect of this is the ability to provide effective information to process operators on the sensitivity of the critical quality attributes (CQAs) of a given product to its critical process parameters (CPPs), and therefore the range of conditions that the process should be operated at in order to produce quality products. This range of conditions can be characterised and communicated effectively via the use of design spaces. It is important to note that changes to the operation of a process that fall within the design space are not considered a change - hence regulatory post approval change processes are not required (ICH, 2009). Ultimately, an accurately identified design space will improve the manufacturing flexibility and simultaneously provide higher confidence about final product quality.

This paper aims to provide a robust framework to characterise design spaces with a high degree of confidence, via regression neural networks and global optimisation concepts. Neural networks provide a link between the CPPs and CQAs which vastly improves the computational speed of the optimisation step of the approach presented, in comparison to the sole use of a mechanistic model. The design spaces created in this paper aim to effectively capture the uncertainty in real processes and are therefore highly applicable to industrial manufacturing in the wake of QbD.

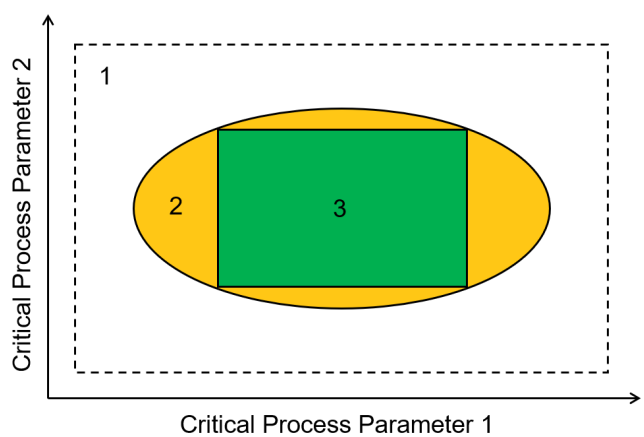


Figure 1.1: *Qualitative plot of the characterised space (1), deterministic design space (2), and operating region (3) for a simple theoretical process with two CPPs*

A simple example of a design space is shown in Figure 1.1. It presents the three crucial regions of a design space: the characterised space, which is the region of all considered operating points; the design space, which is the region in which the target values for the CQAs are satisfied i.e. the region in which 'quality assured operation' (ICH, 2009) will be ensured; and the operating region, a subset of the design space, within which the manufacturing process will be operated at a set point.

The reason for the selection of an operating region (a hyperrectangle) within the design space (rather than use of the entire design space) is due to operational regularity. It is much easier to operate a process within independent, linear ranges in the CPPs (creating an n_θ-orthotopic operating region, where n_θ is the number of CPPs) rather than the use of more complex shapes such as ellipsoids, which will have dependent, non-linear operating ranges. The design space is often highly topologically complex, so the creation of the operating region is advantageous.

The design space construction described in this paper first requires determination of the CPPs and CQAs of a given process. There may be many design variables and measurable outputs for a process, so the choice of the most important CPPs and CQAs is key for the creation of practical design spaces (Maguire and Peng, 2015). Examples of CPPs include temperature, batch time, catalyst loading and oxygen concentration, while examples of CQAs include product purity and minimum production rate. An *in silico* model of the process is also required for the methods presented in this paper, which must be able to model uncertainty in measured parameters. This model needs to be able to provide a continuous mathematical relation between the CPPs and the CQAs.

In this paper probabilistic design spaces are considered, which are design spaces in which every operating point (combination of CPPs) is associated with a probability of reaching the CQA targets, as shown in Figure 1.2. This is opposed to a deterministic design space, which splits the operating points into either a satisfactory (reaches the CQA targets) or unsatisfactory (does not reach the CQA targets) region.

A probabilistic design space presents a key advantage over a deterministic design space, as it is able to capture uncertainty in the measured parameters of the mechanistic model, while a deterministic design space inherently assumes no uncertainty. This is of importance in pharmaceutical systems engineering as the mechanistic or empirical model is built from experimental data, and uncertainty in the kinetic parameters can be significant. Probabilistic design spaces incorporate this uncertainty and are therefore advantageous in accurately characterising the design space.

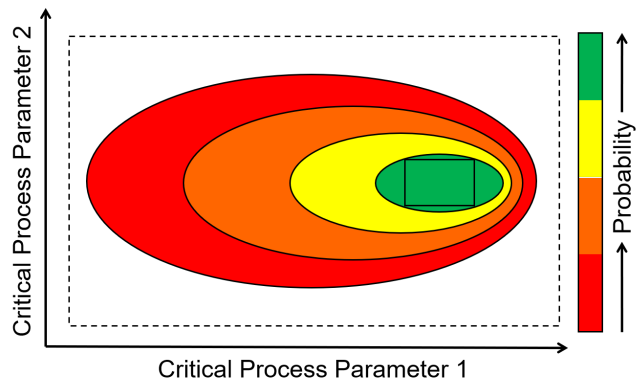


Figure 1.2: *Qualitative plot of the characterised space,* probabilistic *design space, and operating region for a simple theoretical process with two CPPs*

2 Background

To enhance the assurance of safe, effective drug production as called for by Q8, greater system understanding is required in the pharmaceutical industry. This has led to research interest in helping gain further insight into the complex relationships between materials and processes involved in the manufacture of pharmaceutical products. Design spaces are an area of particular research interest, due to the visual understanding they provide as well as the ease of application to industry.

The concept of a design space has been around well before the introduction of QbD. Design spaces share similarities with flexibility analysis from the chemical process industry. Flexibility analysis aimed not just to optimise plant economics, but also considered the flexibility of the manufacturing process (Swaney and Grossmann, 1985). It used multi-level optimisation methods to assess if operation and product quality constraints were satisfied over the entirety of given operating regions. The flexibility index provided a measure of the size of the region of feasible operation, and is closely related to the operating region described in a design space.

'Operational envelopes' which describe regions of operation, rather than operating at fixed set points (Samsatli, Papageorgiou, and Shah, 1999), are one such example of the introduction of the concept of design spaces and hyperrectangles. Multi-level optimisation was initially used to build hyperrectangles based on the assumption that the worst case-scenario would occur at the vertices (Halemane and Grossmann, 1983). Later, more powerful optimisation techniques could identify points that were not necessarily vertices and guaranteed global optimality via relaxation methods (Floudas and Gümüş, 2001).

With the introduction of QbD in the pharmaceutical industry, more modern methods have been used to construct design spaces. Statistical approaches such as Bayesian methods (Peterson, 2008) and multivariate techniques such as Principle Component Analysis and Projection onto Latent Structures (Huang et al., 2009) have been used to construct design spaces, which make efficient use of empirical data. There has been limited use of regression neural networks to approximate mechanistic models in order to create probabilistic design spaces, meaning that this paper is a novel approach to characterising design spaces.

3 Methods

3.1 General Methodology

3.1.1 Data Collection

In order to gather data for the probabilistic characterised space, Monte Carlo simulation was used. For every sampled operating point within the characterised space, N_s Monte Carlo simulations were performed on the mechanistic model using multivariate normally distributed samples in the uncertain reaction parameters, in order to account for measurement uncertainty. It was therefore possible to calculate an approximate probability of success for each operating point, giving co-ordinates in the form $[\theta_1, \theta_2, \ldots, \theta_{n_\theta}, P_{feas}]$, where n_θ is the number of CPPs (denoted as θ) describing the characterisation space, and P_{feas} is the approximate probability that the target CQAs are met for corresponding values in the CPPs. P_{feas} is calculated as N_{feas} / N_s, where N_{feas} is the number of Monte Carlo simulations that meet target CQA values, at corresponding values in the CPPs. These simulations were then repeated for a large range of values in the CPPs, in order to provide sufficient data for the next step of the methodology, effectively defining the characterisation space.

3.1.2 Neural Network Regression

In order to create a continuous design space, it is important that the characterisation space be continuously defined, rather than being a collection of discrete points - hence artificial neural network regression was used to define a continuous algebraic relationship between input and output parameters.

The reason for the use of regression models, rather than the original mechanistic model, is due to the comparative ease of computation. If the mechanistic model were to be used in the optimisation algorithm, then Monte Carlo simulation on a system of ODEs (or even potentially a system of DAEs) would be imposed as a constraint on each point in the characterisation space. This would be much more complex to solve computationally, and may even be infeasible. The neural network provides a single algebraic relation between the CPPs and P_{feas}, making it a preferable alternative.

Neural networks were chosen for use over traditional regression techniques for two main reasons. Firstly, a neural network is suitable for use as there is a large amount of data for it to be trained on - the mechanistic model can be simulated many times without incurring any expenses on the user (except for time). Secondly, neural networks tend to outperform traditional regression techniques when modelling complex relationships between input and output variables, as they are able to model non-linearities, producing more accurate results.

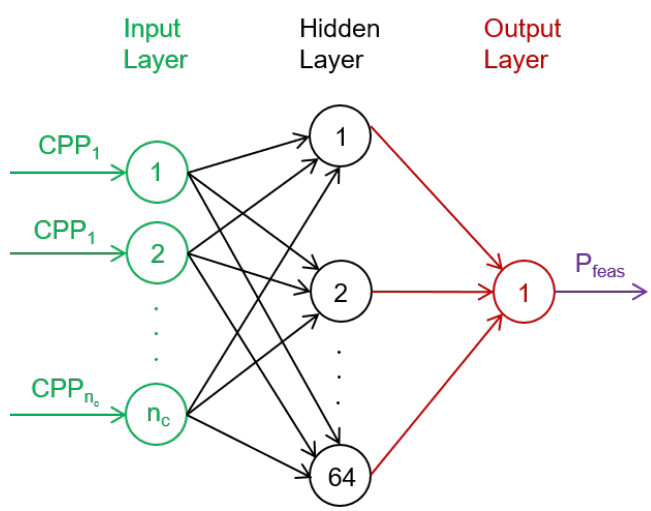


Figure 3.1: *The layout of the neural network used for regression*

The network shown in Figure 3.1 was created using the TensorFlow and Keras libraries in Python, utilising a single 64-neuron hidden layer with sigmoidal activation functions on each neuron, as well as linear activation functions on the input and output layers. Despite the simplicity of the neural network, it has been proven that a feed-forward network with a single hidden layer with sigmoidal activation functions (amongst others) is able to approximate any continuous function (Cybenko, 1989). A hidden layer containing 64 neurons was found to be suitable for both of the examples tested.

The neural network was trained on 100% of the output data from the mechanistic model, as only interpolation was required. In order to ensure that all CPPs are equally weighted by the neural network algorithm, feature scaling (specifically standardisation) was applied to all input data, as the CPPs are often of various orders of magnitude difference. If the data were not feature scaled, then larger inputs would dominate the neural network, and adversely affect the accuracy of prediction.

The network was trained for a maximum of 4000 epochs, with an 'early stopping' mechanism in place if the mean square error of the prediction did not improve for 500 epochs. In order to ensure that the model was able to predict $P_{feas} \pm 0.01$, the mean absolute error and mean standard error of the neural network were measured. As shown in Table 3.1, the mean absolute error for both examples was around 0.01, making the neural networks suitable predictors for the characterised space.

Table 3.1: *Mean Absolute Error (MAE) and Mean Square Error (MSE) for training data in each of the example processes tested*

Error Type	**Example Process**	
	Simple Batch	Suzuki
Mean square	2.585×10^{-4}	2.481×10^{-4}
Mean absolute	0.0102	0.0107

3.1.3 Operating Region Optimisation

With the characterised space continuously defined, it was then possible to define the design space, via setting a minimum bound on $P_{feas} = \alpha$, which corresponds to a (strict) subset of the characterisation space, and contains the collection of points at which $P_{feas} \geq \alpha$. In reality, α is the minimum probability that the process will meet target CQA values that are acceptable to process operators.

The next task was to inscribe the operating region - an n_θ-dimensional hyperrectangle - in the design space. In order to produce the largest operating region possible, optimisation techniques were used to maximise the volume of the corresponding hyperrectangle. The problem can be formulated as a semi-infinite non-linear program as follows:

$$\begin{aligned} \max_{\substack{\theta_1^L,\dots,\theta_{n_\theta}^L \\ \theta_1^U,\dots,\theta_{n_\theta}^U}} \quad & \prod_{i=1}^{n_\theta} (\theta_i^U - \theta_i^L) \\ \text{subject to:} \quad & P_{feas}(\theta_1, \dots, \theta_i, \dots, \theta_{n_\theta}) \geq \alpha, \\ & \forall \theta_i \in [\theta_i^L, \theta_i^U], \forall i = 1, \dots, n_\theta \end{aligned} \tag{3.1}$$

Here θ_i^L and θ_i^U refer to the lower and upper bounds on each CPP in the operating region respectively, and $P_{feas}(\dots)$ is the predictive neural network function within the operating region. The problem with this formulation of the optimisation problem is that the objective function is infinitely constrained - for every point inside the operating region, the probability must be greater than α. As a result, it was not possible to solve this problem computationally. To combat this, a relaxation of the initial problem was formulated, which took the shape of a bilevel optimisation problem. The upper level of this problem can be expressed as:

$$\begin{aligned} \max_{\substack{\theta_1^L,\dots,\theta_{n_\theta}^L \\ \theta_1^U,\dots,\theta_{n_\theta}^U}} \quad & \prod_{i=1}^{n_\theta} (\theta_i^U - \theta_i^L) \\ \text{subject to:} \quad & P_{min} \geq \alpha \end{aligned} \tag{3.2}$$

P_{min} arises as the solution of the lower level problem, which is written as follows:

$$\begin{aligned} \min_{\theta_1,\dots,\theta_{n_\theta}} \quad & P_{feas}(\theta_1, \dots, \theta_i, \dots, \theta_{n_\theta}) \\ \text{subject to:} \quad & \theta_i^L \leq \theta_i \leq \theta_i^U, \forall i = 1, \dots, n_\theta \end{aligned} \tag{3.3}$$

This problem attempts to maximise the volume of the operating region by constraining the minimum probability within the operating region to be greater than or equal to α, as this is sufficient to ensure that every point inside the operating region is feasible.

To solve this bilevel optimisation problem numerically, the extended mathematical programming tool JAMS was initially considered. JAMS is a solver within the GAMS software, which is able to reformulate bilevel NLP problems into Mathematical Programs with Complementarity Constraints

(MPCCs), and then solve them using conventional NLP solvers. It does this by replacing the lower level problem with its respective first-order optimality (KKT) conditions (as equality constraints), and then solving the upper level problem (GAMS Development Corp., 2019).

However, this methodology proved unsuitable for the bilevel problem in question, as first-order KKT conditions assume convexity in the lower level problem (GAMS Development Corp., 2019) in order for the solution to be global. The objective function of the lower level problem is inherently non-convex due to the sigmoidal activation functions, and therefore JAMS was unable to correctly solve the bilevel problem globally.

As the reformulation proved ineffective, another solution was posed to solve the original infinitely constrained optimisation problem, which took the form of a 'generalised cutting-plane algorithm' (Blankenship and Falk, 1976). The procedure required slight reformulation of the constraints of the original problem:

$$\begin{aligned} \max_{\substack{\theta_1^L,\dots,\theta_{n_\theta}^L \\ \theta_1^U,\dots,\theta_{n_\theta}^U}} \quad & \prod_{i=1}^{n_\theta} (\theta_i^U - \theta_i^L) \\ \text{subject to:} \quad & P_{feas}(\theta_1, \dots, \theta_i, \dots, \theta_{n_\theta}) \geq \alpha \\ & \theta_i = \theta_i^L + x_i(\theta_i^U - \theta_i^L), \forall i = 1, \dots, n_\theta \\ & x_i \in [0, 1], \forall i = 1, \dots, n_\theta \end{aligned} \tag{3.4}$$

This reformulation introduces the scaling tensor $\mathbf{x} = [(x_1,\dots,x_{n_\theta})_1, \dots, (x_1,\dots,x_{n_\theta})_{n_c}]$, which relates the values of θ_i to their respective bounds, for each constraint. To then solve the problem, the following iterative procedure was used:

1. Generate an initial single sample for $\mathbf{x}$ in the unit hyperrectangle, $[0,1]^{n_\theta}$, forming the initial constraint ($n_c = 1$)
2. Solve the initial problem with one constraint for each sample point - this took the form of:

$$\begin{aligned} \max_{\substack{\theta_1^L,\dots,\theta_{n_\theta}^L \\ \theta_1^U,\dots,\theta_{n_\theta}^U}} \quad & \prod_{i=1}^{n_\theta} (\theta_i^U - \theta_i^L) \\ \text{subject to:} \quad & P_{k,feas}(\theta_1, \dots, \theta_i, \dots, \theta_{n_\theta}) \geq \alpha, \\ & \forall k \in 1, \dots, n_c \\ & \theta_i = \theta_i^L + x_{i,k}(\theta_i^U - \theta_i^L), \\ & \forall i = 1, \dots, n_\theta, \forall k = 1, \dots, n_c \end{aligned} \tag{3.5}$$

Here, n_c refers to the number of constraints in the current iteration of the algorithm.

3. At the solution of this problem (i.e. the optimal values for θ_i^L and θ_i^U), solve a second optimisation problem of the form:

$$\begin{aligned} \min_{x_1,\dots,x_{n_\theta}} \quad & P_{feas}(\theta_1, \dots, \theta_i, \dots, \theta_{n_\theta}) \\ \text{subject to:} \quad & \theta_i = \theta_i^L + x_i(\theta_i^U - \theta_i^L), \\ & x_i \in [0, 1], \forall i = 1, \dots, n_\theta \end{aligned} \tag{3.6}$$

4. Check if the solution to the previous problem ($= P_{min}$) $\geq \alpha$. If $P_{min} \leq \alpha$, **the corresponding values of x_i are added to the list of samples used as constraints**, and steps 2-4 are repeated. If $P_{min} \geq \alpha$, then the values of θ_i^L and θ_i^U bound the largest possible hyperrectangle, and the original problem is solved.

This process was applied to two separate examples in order to test its efficacy. Firstly, it was tested on a simple theoretical batch reactor with $n_\theta = 2$, as it allowed for effective visualisation of results. It was then tested on a more complex example of a Suzuki coupling reaction, with $n_\theta = 4$. Both of the optimisation problems specified in the algorithm were solved in GAMS using BARON, as it is the most robust non-convex global optimisation solver currently available (Neumaier et al., 2005).

It was also possible to account for the ease of control of each of the CPPs by imposing additional constraints on their minimum allowable ranges. Mathematically, this can be described as $\theta_i^U - \theta_i^L \geq A_i$, where A_i denotes the minumum allowable range in the respective CPP i.

3.2 Simple Process Example

The simple process constituted a single batch reactor intended to optimise the production of a theoretical API, denoted as B. The kinetics of the reactions were assumed to be straightforward:

$$2\,A \xrightarrow{r_1} B \xrightarrow{r_2} C$$
$$r_1 = k_1[A]^2,\ k_1 = A_1 \exp\left(-\frac{E_1}{RT}\right)$$
$$r_2 = k_2[B],\ k_2 = A_2 \exp\left(-\frac{E_2}{RT}\right)$$

The CPPs for the process were the temperature (T) (assuming isothermal operation throughout each batch) and the batch time (τ) of the process. The uncertain parameters for the process were the pre-exponential constants and activation energies for

the two reactions, ($k_{1,2}$ and $E_{1,2}$), and these were normally sampled 200 times for 4,243 (τ, T) coordinate pairs to give the respective P_{feas}, via the related mechanistic model.

3.3 Suzuki Coupling Process Example

The Suzuki coupling reaction was more complex than the simple batch process - while still a batch reactor, the system was in vapour-liquid equilibrium, and contained 17 chemical species. The active species participated in 12 chemical reactions, of which three were reversible, and one was considered to be instantaneous. The four corresponding CPPs for this process were the batch time, the catalyst loading, the temperature, and the molar fraction of O_2 in the reactor head space. The uncertain reaction parameters were the fourteen pre-exponential factors of the kinetically limited reactions, and these were normally sampled 1,000 times for each of the 13,647 data points in the characterisation space. This data was provided by the Centre for Process Systems Engineering at Imperial College London.

4 Results and Discussion

4.1 Simple Batch Reactor

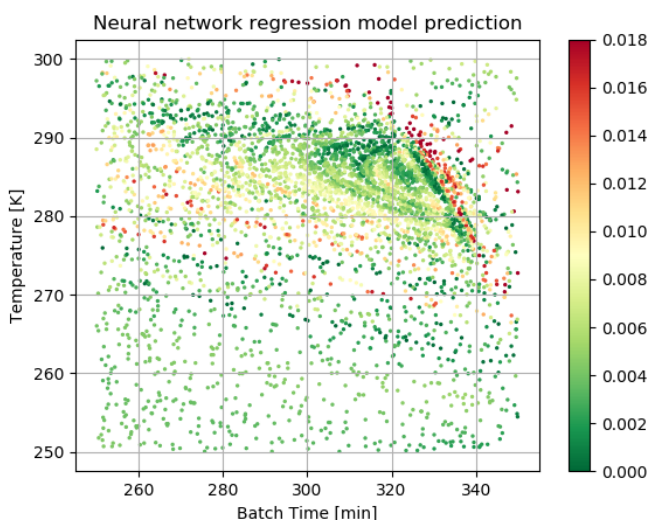


Figure 4.1: *Comparison of neural network prediction and model data. The colour map represents the absolute error in prediction of P_{feas}*

As presented in Table 3.1, regression neural networks are an excellent predictor for the mechanistic model in this example. Figure 4.1 helps to further visualise this, and shows that the regression model is accurate for both the extrema of the design space and intermediate points, with a maximum absolute error of prediction of P_{feas} of around 0.018.

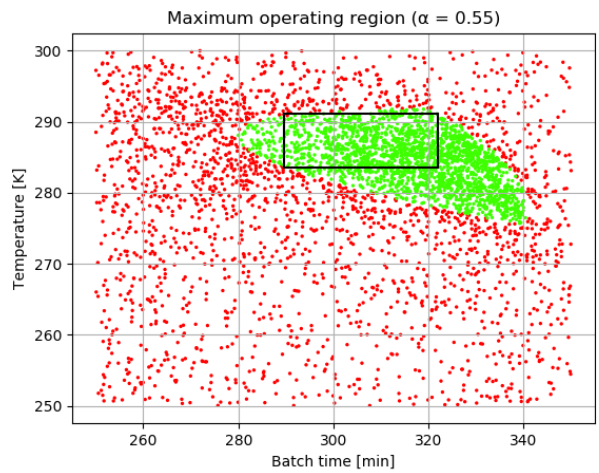


Figure 4.2: *Operating region defined by optimisation compared to original mechanistic model data. Feasible points are in green, with infeasible points in red*

Figure 4.2 shows that the algorithmic approach used was able to find the position and size of the largest possible operating region within the probabilistic design space. It is observed that not every point inside the operating region is feasible - in the top left corner of the operational envelope, there is an infeasible point. This is likely due to the small mean average error in the neural network prediction, and could be combated by increasing the value for α by a small amount, or further refining the neural network. Despite this small discrepancy, Figure 4.2 shows that the approach is accurate and viable in this scenario. The corresponding values of batch time and temperature are enumerated in Table 4.1. It is also worth noting that here, α was selected to be 0.55, as the maximum feasibility of this process was only around 0.75 - in pharmaceutical contexts, α would be expected to be higher.

Table 4.1: *Upper and Lower bounds on CPPs defining operating region at α = 0.55 for the simple batch process*

Critical Process Parameter	θ_i^L	θ_i^U
Batch Time [min]	289.5	322.0
Temperature [K]	283.5	291.2

It is worth noting that three of the four corners of the inscribed operating region lie on the boundary between satisfactory and unsatisfactory points, reducing flexibility at these extrema. This could be accounted for by increasing α slightly above the necessary level, but it is more important to note that

normal operation would not usually occur at these extrema, but in the centre of the operating region.

It was also possible to visualise the effect of adding additional control constraints on the CPPs. In Figure 4.3, a minimum temperature range of 10K was imposed on the optimisation process, altering the size and position of the operating region, without negatively affecting accuracy.

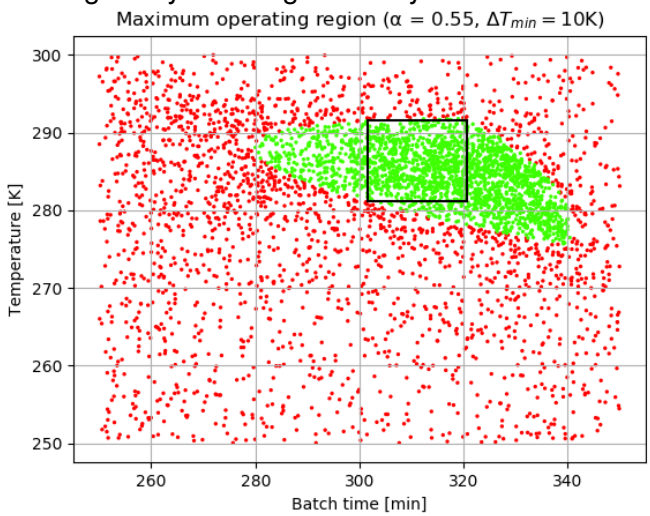


Figure 4.3: *Maximum operating region with a minimum temperature bound of 10K, compared to model data*

4.2 Suzuki Coupling Process

The four CPPs in the Suzuki coupling example correspond to an operating region in a four-dimensional design space. The optimisation algorithm successfully found the largest operating region within the design space which is in the shape of a hyperrectangle in four dimensions (4-orthotope). Here, α is 0.95, meaning every point within the hyperrectangle corresponds to a minimum of 95% probability of reaching the CQA targets. Table 4.2 shows the results of the optimisation approach in this four dimensional example.

Table 4.2: *Upper and lower bounds on CPPs defining operating region at α = 0.95 for the Suzuki coupling process*

CPP	θ_i^L	θ_i^U
Batch Time [min]	109.5	185.1
Temperature [K]	318.9	338.8
Catalyst Loading [-]	0.0017	0.0029
O_2 Mole Fraction [-]	2.43×10^{-6}	2.55×10^{-4}

The results were validated by extensive Sobol sampling within the hyperrectangle using the neural network to confirm that the minimum probability was always satisfied. Over 2 million sobol samples were taken in this validation.

The four-dimensional operational envelope is difficult to show in a plot. For visual purposes, one CPP was fixed at its mean value to create three dimensional plots. Figure 4.4 shows a 3-D scatter plot of the design space and operating region with varying temperature, batch time and catalyst loading with the oxygen mole fraction fixed at its mean value from mechanistic data. (1.29×10^{-4}).

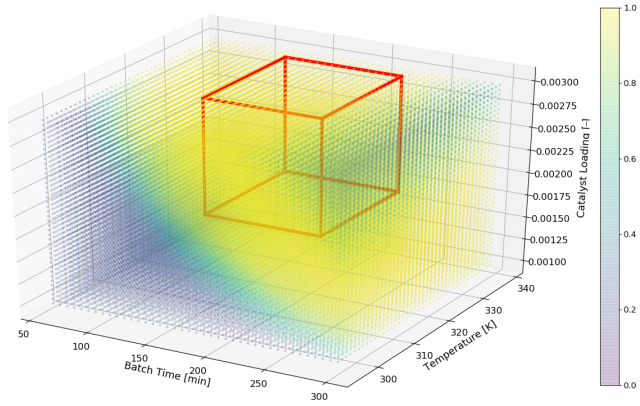
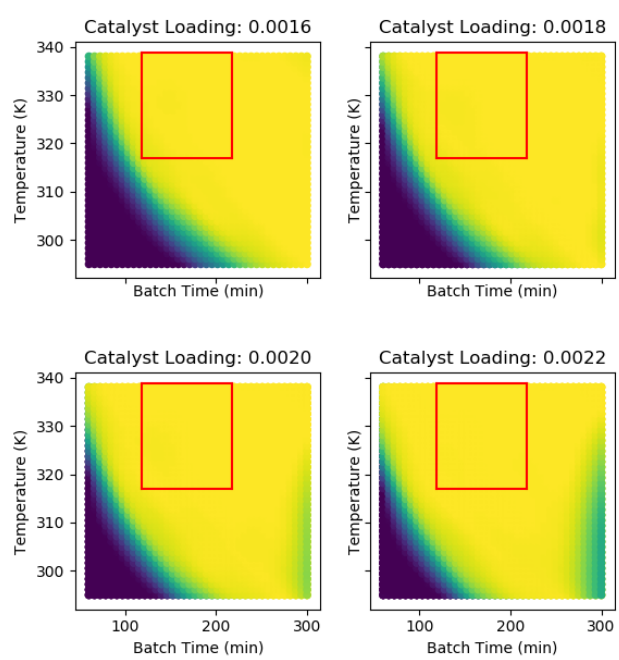


Figure 4.4: *Operating region (in red) at α = 0.95 with varying batch time, temperature and catalyst loading. Colour gradient represents P_{feas}*

An alternative way of presenting the data is through trellis plots, which provide more effective visualisation of 3-dimensional data, allowing for simple data validation. Figure 4.5 presents the same data as shown in Figure 4.4. Each subplot slices the design space into a 2D plot at a given value for catalyst loading.



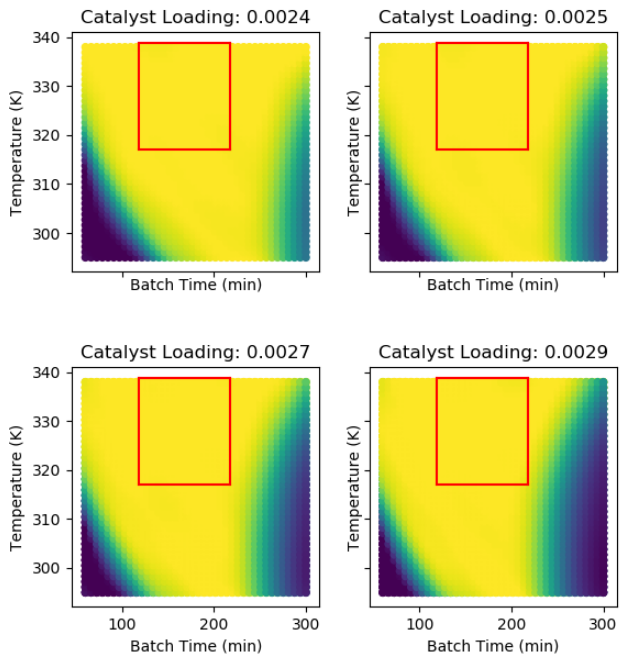


Figure 4.5: *Trellis sub-plots showing operating region (in red) at α=0.95 with varying batch time, temperature and catalyst loading. Colour gradient represents* P_{feas}

5 Conclusion

A novel, robust machine-learning oriented approach has been outlined for characterising probabilistic design spaces and operating regions. This approach requires a mathematical model of the process, which must incorporate the effects of uncertainty in measurable parameters, such as Arrhenius constants. If the model is unable to account for uncertainty, then it is not possible to generated a probabilistic design space, and the neural network regression techniques employed will not be able to effectively characterise the (instead) deterministic design space. The use of experimental data could also be a viable option for the procedure, although a large number of experiments must be run in order to collect enough data to accurately train the neural network to account for uncertainty.

Given a model of a process, a regression neural network is an effective, accurate method for providing a continuous definition of a probabilistic design space, and is readily able to be further utilised in the formation of the largest possible operating region within the design space. The resulting operating region has been shown to be accurate to the initial input data, meaning that the solution approach of using machine learning is viable in this context.

Through use of machine learning to characterise the probabilistic design space, it is possible to employ global optimisation techniques in order to inscribe the largest possible operating region at a given minimum acceptable P_{feas}. The operating region is of benefit to process operators as it can be described with continuous, independent ranges in the CPPs, and does not require visualisation of the data. This is of especial importance in pharmaceutical contexts as there are often many CPPs, meaning that the resulting design space consists of more than three dimensions, hindering data visualisation.

The work done in this paper also highlights the importance of the development of solvers designed for bilevel optimisation problems. Due to their reformulation strategies, current solver packages are unable to correctly solve bilevel problems in which the lower-level problem exhibits non-convexity, meaning that a globally optimum solution to the original problem cannot be found. If a global bilevel solver were to become commercially available, then there may be an opportunity to further expedite the approach used.

6 Outlook

Despite the success of the approach taken, further improvements may be possible through adjustment of certain steps of the solution procedure. One of the ways in which the regression neural network could be improved is via adjustment to the number of neurons in each hidden layer of the network, as well as the number of hidden layers, and the activation function used. While the mean average error of the network was shown to be within acceptable bounds in both cases, it may be possible to further reduce this via adjustment to the aforementioned parameters. These adjustments may also help to reduce the amount of training time required for the network.

It may also be possible to improve the optimisation performed in GAMS. Additional initial constraints could be imposed in the problem formulation. By doing this, it may reduce the number of times the system would need to loop in order to find the largest possible operating region, which may save time.

Neural networks may also be used to classify deterministic design spaces in pharmaceutical processes. Rather than performing Monte Carlo sim-

ulation to generate a probabilistic design space, a single simulation with no uncertainty can be performed at each point, and then classification neural networks can be used to characterise the deterministic design space.

7 Potential Applications

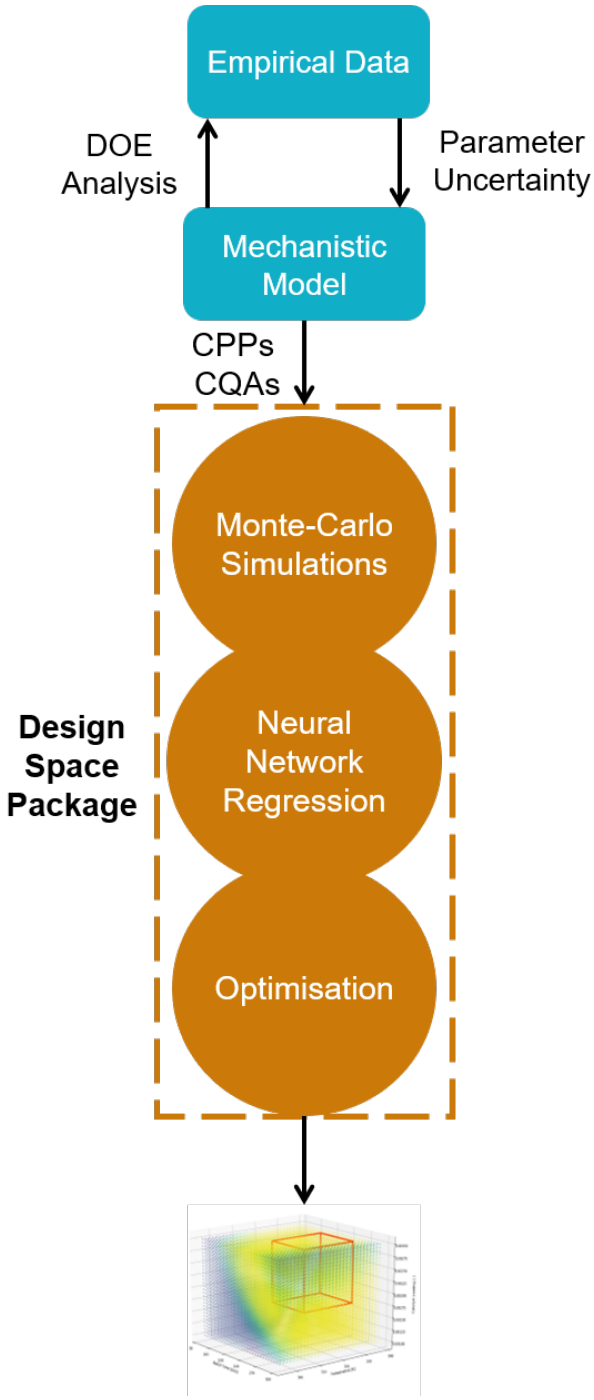


Figure 7.1: *Simple flowchart of potential optimisation package*

The introduction of QbD has highlighted the importance of design spaces in helping to gain a deeper understanding of pharmaceutical processes. However, the pharmaceutical industry has been slow to adopt new methods of design space characterisation, often sticking to empirical and deterministic definitions. To help with this, the methods presented in this paper could be combined into a compact, user-friendly package.

This package could then be used by process operators and researchers to help construct design spaces for their own purposes. The only requirement for the design space package would be a mathematical model to describe the process, which must be able to either directly or indirectly relate the CPPs to the CQAs. The model must be able to account for parameter uncertainty, in order to create the probabilistic design spaces. An outline of how the package would look is shown in Figure 7.1.

8 Acknowledgements

The authors would like to express their appreciation to Kennedy Kusumo, a PhD student at Imperial College, for his guidance and support throughout the duration of this project, and his willingness to dedicate time towards helping troubleshoot the issues that occurred.

References

Blankenship, J. W. and J. E. Falk (June 1976). "Infinitely constrained optimization problems". In: *Journal of Optimization Theory and Applications* 19.2, pp. 261–281. ISSN: 1573-2878. DOI: 10.1007/BF00934096. URL: https://doi.org/10.1007/BF00934096.

Cybenko, George (1989). "Approximation by Superpositions of a Sigmoidal Function". In: *Math. Control Signals Systems* 2, p. 303. DOI: https://doi.org/10.1007/BF02551274. URL: https://link.springer.com/article/10.1007/BF02551274.

Floudas, Christodoulos A. and Zeynep H. Gümüş (2001). "Global Optimization in Design under Uncertainty: Feasibility Test and Flexibility Index Problems". In: *Industrial & Engineering Chemistry Research* 40.20, pp. 4267–4282. DOI: 10.1021/ie001014g. URL: https://doi.org/10.1021/ie001014g.

GAMS Development Corp. (2019). *JAMS and LogMIP*. URL: https://www.gams.com/latest/docs/S_JAMS.html#EMP_BILEVEL_PROGRAMS (visited on 04/12/2019).

Halemane, K. P. and I. E. Grossmann (1983). "Optimal process design under uncertainty". In: *AIChE Journal* 29.3, pp. 425–433. DOI: 10.1002/aic.690290312. URL: https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.690290312.

Huang, J. et al. (Dec. 2009). "Quality by design case study: an integrated multivariate approach to drug product and process development". In: *Int J Pharm* 382.1-2, pp. 23–32. DOI: 10.1016/j.ijpharm.2009.07.031.

ICH (2009). *Pharmaceutical Development Q8(R2)*. URL: https://database.ich.org/sites/default/files/Q8_R2_Guideline.pdf.

Maguire, Jennifer and Daniel Peng (2015). "How to Identify Critical Quality Attributes and Critical Process Parameters". In: (North Bethesda, Maryland). US Food and Drug Administration. URL: http://pqri.org/wp-content/uploads/2015/10/01-How-to-identify-CQA-CPP-CMA-Final.pdf.

Neumaier, Arnold et al. (June 2005). "A Comparison of Complete Global Optimization Solvers". In: *Math. Program.* 103, pp. 335–356. DOI: 10.1007/s10107-005-0585-4.

Peterson, J. J. (2008). "A Bayesian approach to the ICH Q8 definition of design space". In: *J Biopharm Stat* 18.5, pp. 959–975. DOI: 10.1080/10543400802278197.

Samsatli, N.J., L.G. Papageorgiou, and N. Shah (1999). "Batch process design and operation using operational envelopes". In: *Computers & Chemical Engineering* 23. European Symposium on Computer Aided Process Engineering, S887–S890. ISSN: 0098-1354. DOI: https://doi.org/10.1016/S0098-1354(99)80218-X. URL: http://www.sciencedirect.com/science/article/pii/S009813549980218X.

Swaney, R. E. and I. E. Grossmann (1985). "An index for operational flexibility in chemical process design. Part I: Formulation and theory". In: *AIChE Journal* 31.4, pp. 621–630. DOI: 10.1002/aic.690310412. eprint: https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.690310412. URL: https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.690310412.

Techno-Economic Feasibility Assessment of Oxyfuel Combustion Applied in the Cement Industry

Tumisang Tsolope and Nonofo Thapelo
Department of Chemical Engineering, Imperial College London, U.K.

Abstract: This paper presents an assessment of technical and economic implications of operating a full oxyfuel cement plant as modelled using Aspen Plus® v9. Using oxyfuel combustion technology led to a higher production of clinker compared to the base plant without Carbon Capture and Storage (CCS) and did not compromise the quality of the clinker. The oxyfuel cement plant captured 98.9% of the CO_2 produced with a cost of CO_2 avoided of €50.81 per tonne of CO_2 (t_{CO2}). The cost of production increased from €51.22 per tonne of clinker (t_{clk}) in the base plant without capture to €80.15/t_{clk} for the oxyfuel cement plant. The increase in the cost of production was mainly due to increase in the electricity usage and additional capital invested in the air separation unit and CO_2 processing unit. Electricity usage increased by 80% in the oxyfuel plant which was partially mitigated by the use the use of an Organic Rankine Cycle to produce additional electricity from waste heat. This study in one few to consider viability of heat recovery through the use of an ORC on a oxyfuel cement plant. A flue gas recirculation ratio of 70% was found to be the optimum to achieve similar operation of the oxyfuel plant as that of the base plant without capture. This indicates that the adoption of oxyfuel technology is a viable option to mitigate CO_2 emissions in the cement industry but further research and development is necessary to reduce power consumption and understand the technological complexity of the process.

Keywords; Oxyfuel; Aspen Plus®; Clinker; Cost of Avoided CO_2; Air Separation Unit (ASU); CO_2 Processing Unit (CPU); Organic Rankine Cycle (ORC); Flue gas recirculation (FGR);

1. Introduction

Global cement production was estimated to be 4.65 billion tonnes in 2017 (CEMBUREAU, 2017). This accounted for nearly 8% of the total direct carbon dioxide (CO_2) emissions globally (BBC, 2018). The annual global production is expected to grow by 3.5% per annum from 2019-2024 (IMARC Group, 2018). During cement production, the CO_2 emissions come not only from the combustion of fossil fuels but also from the calcination reaction which accounts for 60% of the total CO_2 produced. The combustion of fossil fuels for calcination and clinker production accounts for the remaining 40% of CO_2 produced. Other indirect CO_2 comes from electricity production and transportation (Carrasco-Maldonado et al., 2015).

The key strategies that can be employed to reduce the emissions from cement production include improving energy efficiency, reducing the clinker/cement ratio by using more additives in cement, using low carbon fuels and carbon neutral fuels (IEA, 2009). However, these methods are limited because they do not tackle the CO_2 associated with the calcination reaction. The implementation of Carbon Capture and Storage (CCS) technologies into cement plants has been suggested by the IEA *Cement Technology Roadmap* as a significant tool to reduce the CO_2 emissions associated with the calcination reaction (IEA, 2009).

There are a few existing technologies for CCS: pre-combustion, post combustion and oxyfuel combustion. Pre-combustion methods have not been widely taken up because the majority of emissions, those from the calcination reaction, would remain unabated if this method were to be adopted. Post-combustion technologies are end of pipe methods that would not require significant retrofit to existing plants (Cormos & Cormos, 2017). Examples of end of pipe methods include calcium looping and chemical adsorption using amines and other solutions (IEA, 2009). Oxyfuel combustion uses oxygen instead of air for combustion in a cement kiln. This eliminates the presence of nitrogen (N_2) in the kiln which results in a flue gas stream that mainly consists of CO_2 and water (H_2O) (Carrasco-Maldonado et al., 2015). In oxyfuel combustion, flame temperatures can be excess of 3500°C, which would cause serious damage to the kiln. Thus, a proportion of the CO_2 rich flue gas is recycled back to the combustion zone to dilute the oxygen rich stream to lower the temperature of the kiln (Barker et al., 2009).

This paper describes the technical and economic analysis of employing oxyfuel combustion in modern cement plants. The objectives of this research are as follows:

- Model a baseline European cemen plant without capture and an oxyfuel cement plant in Aspen Plus producing approximately 1 Mt of cement per year.
- Investigate the capture efficiency of an oxyfuel cement plant.
- Investigate the effect of oxyfuel combustion on clinker quality.
- Perform an economic analysis of the baseline and oxyfuel cement plants.

2. Process descriptions

2.1 Base plant

2.1.1 Preheater

The preheater usually consists of three to five cyclone stages with one above the other. Hot gases from the pre-calciner enter the preheater at a temperature of approximately 900°C. This results in 20% calcination of the raw meal in the preheater. The flue gases leave the top stage of the preheater at 330°C (Barker et al., 2009).

2.1.2 Pre-Calciner

Secondary fuel combustion is conducted in the pre-calciner. About 60% of the total fuel is fed into the pre-calciner to provide heat for the calcination of raw meal. The raw meal leaves the pre-calciner at 870°C and is 75-95% calcined (European Commission, 2001). Hot gases from the kiln enter the pre-calciner at 1000°C and leave at 900°C (Barker et al., 2009).

2.1.3 Kiln

The kiln operates at temperatures of 1400-1500°C (Winter, 2012). In total, 40% of the fuel is fed to the kiln to provide the energy to attain the high temperatures. The calcined raw meal from the pre-calciner enters the kiln at 870°C where it reacts with silica, alumina and ferrous oxide to form the major four phases of clinker i.e. alite, belite, tricalcium aluminate and tetracalcium alumino-ferrite. (Rahman et al., 2017).

2.1.4 Clinker cooler

The clinker cooler has two main functions; to reduce the temperature of hot clinker (1450°C) for downstream process equipment and to recover as much heat as possible for use in the process. The heat is recovered by preheating the air used for the main and secondary combustion (Schorcht et al., 2013).

2.2 Oxyfuel plant additions

2.2.1 Air separation unit (ASU)

The ASU is required to separate oxygen from air to provide a pure oxygen stream for oxyfuel combustion. Cryogenic oxygen production is currently the most mature technology (Laribi et al., 2017). The cryogenic air separation unit has a typical power consumption of 200-240 kWh/tO_2 (Barker et al., 2009). For a typical cement plant with a kiln capacity of 3000t/d, the oxygen demand is estimated to be 0.25-0.35 t_{O_2}/t_{clk} (Carrasco-Maldonado et al., 2015).

2.2.2 CO_2 Purification Unit (CPU)

The CPU is required to separate CO_2 from the flue gases exiting the preheater. The CPU consists of 3 main units; the sour compression unit (SCU) for removal of NO_x and SO_x, the dehydration unit for the removal of water from the flue gas and a cryogenic unit for liquefaction of CO_2 for storage and transport (Meunier et al., 2014).

3 Background

Research into oxyfuel cement production has attracted a lot of interest as some of the preliminary models show that it has a lower Cost of Avoided CO_2 (CAC) when compared with other methods. Gardarsdottire et al. (2019) has estimated it to have a CAC of €$42.4/t_{CO_2}$. The second cheapest method of CCS was calcium looping whose CAC was estimated to be €$66.2/t_{CO_2}$. Research has found that there is a difference in performance between the partial and full oxyfuel combustion configurations. The partial configuration has two preheaters of which one together with the kiln are air operated while the other preheater is oxygen operated. The full configuration on the other hand uses an air operated kiln and only uses one preheater. Full oxyfuel was found to have a lower CAC than partial oxyfuel (€$43.0/t_{CO_2}$ vs €$55.2/t_{CO_2}$) (IEA, 2013). The full oxyfuel configuration also exhibited higher capture efficiencies of up to 99% compared to the possible 75% by the partial configuration.

The need for efficient treatment of flue gases to remove mercury, nitrogen and sulphur oxides from the flue gas has been highlighted as an important aspect to consider when using oxyfuel combustion due to their toxicity and greenhouse effect (Carrasco-Maldonado et.al. 2016). The ASU and CPU were found to increase the overall energy consumption by 86% and 128% in partial and full oxyfuel combustion respectively (Barker et al., 2009). The first pilot plant on the oxyfuel technology was done in 2009 by FLSmidth, Lafarge and Air Liquide. It was a 2-3tonne/day plant with CO_2 capture at a cost of $€62/t_{CO_2}$. The European Cement Research Academy in conjunction with the International Energy Agency are also in the process of building a fully-fledged plant that is expected to provide more insight on the technology and is expected to start operation in 2025 (IEA, 2013).

4 Methods

The base plant without capture and the oxyfuel cement plant were both modelled on Aspen Plus v9 using the Ideal thermodynamic property method. This is because this model involves solids handling where vapour-liquid equilibrium (VLE) is considered unimportant. The assumptions used to develop the models are shown in Table 1 below.

Table 1 Assumptions used for modelling

Raw meal	**Temperature (°C)**	60
	Composition (wt%)	78% $CaCO_3$, 13.9% SiO_2, 3.3% Al_2O_3, 2% Fe_2O_3, 1.5% $MgCO_3$, 1% H_2O
Fuel	**Temperature (°C)**	60
	Composition (wt%)	Proxanal
		25% Moisture, 45.1% Fixed Carbon, 45.7% Volatile matter, 9.2% Ash
		Ultanal
		9.2% Ash, 67.1% C, 4.8% H, 1.1% N, 0.1% Cl, 1.3% S, 16.4% O
		Sulfanal
		0.6% Pyritic, 0.1% Sulfate, 0.6% Organic
	Heating value (MJ/kg)	27.2
Preheater		
Number of stages	5	
Cyclone efficiency (1st - 5th), %	96/86/86/86/76	
Calciner		
Calcination efficiency	100	
Fuel consumption (kg/s)	2.45	
Kiln		
Fuel consumption (kg/s)	1.51	

4.1 Modelling Principles

4.1.1 Coal Treatment

The coal used for this model is lignite coal. The composition of the coal used in shown in Table 1. As this coal has a high moisture content (25%), it was first dried using nitrogen in an RStoic reactor according to AspenTech's manual for drying solids (Aspen Technology, 2013). No energy was provided for the drying process. A high moisture content in coal delays combustion and makes the kiln flame longer (Chatterjee, 2018). A high coal moisture content also results in an increase in the volume of flue gases due to the evaporation of the moisture. This leads to increased fan power and thus lower efficiency in the plant (Bullinger et al., 2002).

4.1.2 Preheater

The preheater tower was modelled as a five-stage preheater with the cyclone efficiencies shown in Table 1. Each stage of the cyclone is modelled using 2 mixers (A, D) and 2 splitters (B, C) as indicated on Figure 1. In stage 1, the raw meal and the flue gas from the second stage are fed to mixer A where the solid is heated up by the gas. The resulting mixture is fed to splitter B which completely separates the gas phase from the solid phase. The gas is sent to mixer D while the solid is sent to splitter C. Splitter C then splits the solid according to the cyclone efficiency. In mixer D, the solid and gas mix to form flue gas which contains dust while the rest of the solid from splitter C is delivered to the next stage of the preheater.

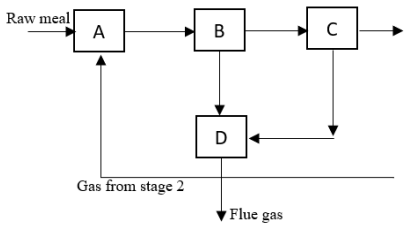


Figure 1 Preheater model

4.1.3 Calciner

The calciner was modelled using 3 reactor units; RYield, RGibbs and RStoich. The RYield reactor was used to model the decomposition of coal while the RGibbs reactor was used to simulate coal combustion. Both these reactors were modelled according to AspenTech's manual (Aspen Technoloy 2013) on modelling processes with solids. The RStoich reactor was used to model the complete (100% conversion) calcination of $CaCO_3$ and $MgCO_3$ according to the following equations:

$$CaCO_3 \rightarrow CaO + CO_2 \quad (1)$$
$$MgCO_3 \rightarrow MgCO + CO_2 \quad (2)$$

The operating temperature of the calciner is assumed to be 870°C while the tertiary air temperature is assumed to be 1100°C. 66% of the fuel used is consumed in the calciner.

4.1.4 Rotary kiln

The kiln was also modelled using 3 reactor units; RYield, RGibbs and RStoich. The RYield and the RStoich were modelled the same way as in the calciner. The RStoich was used to simulate the reactions occurring the kiln to form the major phases of clinker.

For an output of 1Mtonne of clinker per annum, a mass balance using the Bode equations (Winter, 2012) was carried out to calculate the raw meal feed rate. The mass balance was carried out based on the typical composition of clinker as indicated below, where the abbreviations are shown in Table 2.

$$C_3S = 4.0710CaO - 6.024SiO_2 - 1.4297Fe_2O_3 - 6.7187Al_2O_3 \quad (3)$$

$$C_2S = 8.6024SiO_2 + 1.0785Fe_2O_3 + 5.0683Al_2O_3 - 3.0710\ CaO \quad (4)$$

$$C_3A = 2.6504Al_2O_3 - 1.6920Fe_2\ O_3 \quad (5)$$

$$C_4AF = 3.0432Fe_2\ O_3 \quad (6)$$

Table 2 Typical clinker composition (Winter 2012)

Phase	Mass Composition (%)	Chemical Formula (Abbreviation)
Alite	65	Ca_3SiO_5 (C_3S)
Belite	15	Ca_2SiO_4 (C_2S)
Tricalcium Aluminate	7	$Ca_3Al_2O_6$ (C_3A)
Tetracalcium Alumino-ferrite	8	$Ca_4Al_2F_2O_{10}$ (C_4AF)

The calculation of the mass balance can be found in the supplementary materials section 2.

The reactions occurring in the kiln are shown in **Error! Reference source not found.**. Tetracalcium alumino-ferrite (C_4AF) was not available in Aspen databanks hence it was not included in the reaction matrix. It was also

assumed that 75% of C_2S is used in reaction 8 to produce C_3S and that the conversion of Al_2O_3 in reaction 9 is 98% (Guo, 2019). 34% of the coal used in the plant is consumed in the kiln.

Table 3 Reactions occurring in the kiln (Guo, 2019)

Reaction	Reaction Equation	Reaction Enthalpy (kJ/kg)	Reaction Number
Formation of Oxides and Decomposing Reactions			
Evaporation of Water	$H_2O_{(l)} \rightarrow H_2O_{(g)}$	2453	1
Dissociation of Kaolinite	$AS_4H_2 \rightarrow Al_2O_3 + 2SiO_2 + 2H_2O$	748	2
Oxidation of Carbon	$C + O_2 \rightarrow CO_2$	-33 913	3
Dissociation of $MgCO_3$	$MgCO_3 \rightarrow MgO + CO_2$ (g)	1395	4
Dissociation of $CaCO_3$	$CaCO_3 \rightarrow CaO + CO_2$ (g)	1780	5
Formation of Intermediates			
Formation of Belite (C_2S)	$2CaO + SiO_2 \rightarrow C_2S$	-732	6
Formation of SO_2	$S + O_2 \rightarrow SO_2$	-17813	7
Formation of Final Phases			
Formation of Alite (C_3S)	$C_2S + CaO \rightarrow C_3S$	59	8
Formation of Tricalcium Aluminate (C_3A)	$3CaO + \alpha\text{-}Al_2O_3 \rightarrow C_3A$	-27	9
Formation of Calcium Alumino-ferrite (C_4AF)	$4CaO + \alpha\text{-}Al_2O_3 + \alpha\text{-}Fe_2O_3 \rightarrow C_4AF$	-105	10

4.1.5 Clinker Cooler

To simulate the clinker cooler, 3 heat exchangers were used with the clinker as the hot stream and air as the cold stream. The secondary and tertiary air streams were at 637K and 692K respectively. It was assumed that in each heat exchanger, 1% of the clinker is lost as dust.

4.2 Oxyfuel Combustion Specific Units

4.2.1 Air Separation Unit (ASU)

The ASU was not modelled in this project. Instead it was assumed that oxygen is received as 99% pure O_2 with 1% N_2. However, the electricity consumption, CAPEX and OPEX of a typical ASU were taken into consideration when preparing the financial model.

4.2.2 CO_2 Processing Unit (CPU)

The CPU has 3 main components; the sour compression unit (SCU) where the nitric and sulphur oxides in the gas are captured. The dehydration unit which removes water. The cryogenic unit where the carbon dioxide is compressed and liquified. The SCU and the dehydration units were modelled in a single model using the NRTL thermodynamic method. This method was chosen because the liquid mixture was assumed to be non-ideal. The cryogenic unit model used UNIQUAC method as only phase equilibria were considered (Meunier et al. 2014).

4.2.2.1 Sour Compression Unit (SCU)

The SCU was modelled according to previous work by Meunier et al. (2014). Two absorber columns were used to remove the NOx and SOx respectively. They were both modelled using the RadFrac column. The gas is compressed from 1 bar to 15 bars in a two-stage isentropic compressor before being sent to the desulphurisation column. Water is used as the scrubbing liquid on economical grounds. The gas is then compressed to 30 bars before being sent to the denitrification column. In this column 90% of the water leaving this column is recycled and it is mixed with make-up stream before being fed back into the column. Both columns are sized equally as shown in Table 4. These absorbers were designed for gas and liquid flow rates of 3000 Nm /h and 15 m /h respectively leading to a liquid/gas ratio of 0.005.

Table 4 Dimensions of the absorption columns (Meunier et al., 2014)

Dimension	Value
Internal diameter (m)	0.15
Packing height	12m
Number of stages	16
Packing used	IMPT 25

4.2.2.2 Dehydration unit

The washed gas stream was then sent to the dehydration unit. This unit was modelled as an absorber using a RadFrac column. Triethylene Glycol (TEG) was used as the scrubbing liquid. TEG was chosen due to its ease of regeneration (high thermal and chemical stability), moderate price and the high affinity between the water molecules and the hydroxyl and ether groups in TEG (Nivargi, 2005). 70% of the TEG exiting the column was mixed with a make-up steam and recycled back into the column. The column was trayed with diameter 1.13m, height 5.5m.

4.2.2.3 Cryogenic Unit

The cryogenic unit was modelled based on previous work by Meunier et al. (2014). After dehydration the gas was sent to the cryogenic unit where it initially was cooled to -30ºC before being sent to the first flash unit where some liquid carbon dioxide was collected. The vapour stream from the first flash was further cooled to -55ºC. Caution was taken to avoid cooling the gas below -55ºC to avoid formation of solid CO_2 (dry ice). It was then sent to the second flash unit and liquid carbon dioxide was collected again. The streams of liquid CO_2 from the two flash units were mixed and pumped to 110 bars to prepare it for transportation and storage.

4.2.3 Organic Rankine Cycle

The organic Rankine cycle is used to recover heat from the flue gas stream. Process water is used to generate steam for the heat recovery. Two turbines and two heat exchangers for this process. Steam is fed into the turbines at 10 bar and 499K and at 15 bar and 499K. The steam exiting the turbines is saturated at 1atm.

4.3 Financial model

The financial analysis was performed according to methods suggested by Sinnott R.K. in *Chemical Engineering Design (2005)*. The assumptions associated with the financial model are listed in the **Error! Reference source not found.** below.

Table 5 Cost assumptions for the financial model

Assumptions for the cost basis	
Cost basis	€ 2018
Capacity factor (%)	87
Tax rate (%)	0
Operating life (years)	25
Construction time (years)	3
Percentage of TPC depreciated (%)	100
Inflation rate (%)	0
Discount rate (%)	8

The capital and operating costs were obtained from the knowledge of costs from earlier studies by Gerbelová et al. (2016) and IEA (*Deployment of CCS in the Cement Industry,2013*) which use similar clinker manufacturing processes. The studies were used to obtain the prices of the main plant equipment (e.g. kiln, ASU, CPU and fans) and other engineering costs while the utility operating costs were based on the Aspen model that was developed. These utility operating costs were based on current prices of water, electricity and chemicals (i.e. TEG).

The capital cost can be estimated using the following equation if the capacity of a similar unit is known. The value of the index n is usually taken as 0.6; the well-known six-tenths rule.

$$C_2 = C_1 \times \left(\frac{S_2}{S_1}\right)^n \qquad (7)$$

These costs were readjusted to 2018 values to account for the time value of money using equation 8.

$$C_A = C_B \times \frac{CEPCI\ in\ year\ A}{CEPCI\ in\ year\ B} \qquad (8)$$

where C_A and C_B are the costs in year A and year B and CEPCI is the Chemical Engineering Plant Cost Index.

The financial performance of the capture technology is evaluated using the *cost of clinker (COC)* and the *cost of CO_2 avoided (CAC)*.

The *cost of clinker (COC)* is calculated by adding the contributions of the annualized capital cost C_{cap}, the cost of fuel C_{fuel}, the cost of raw materials C_{RM}, the cost of electricity C_{el} and all other operating and maintenance costs $C_{O\&M}$, all expressed as per ton of clinker (i.e. €/t_{clk}).

$$COC = C_{cap} + C_{fuel} + C_{RM} + C_{el} + C_{O\&M} \quad (9)$$

The *cost of CO_2 avoided (CAC)* is evaluated by comparing the cost of clinker and the equivalent specific emissions in the reference plant and the oxyfuel combustion plant. It is expressed as:

$$CAC = \frac{COC - COC_{ref}}{e_{clk,eq,ref} - e_{clk,eq}} \quad (10)$$

where $e_{clk,eq,ref}$ is specific equivalent emissions from the reference kiln, and $e_{clk,eq}$ is the specific equivalent emission from the kiln with capture.

Specific equivalent emissions are expressed as the sum of direct e_{clk} and indirect $e_{el,clk}$ emissions:

$$e_{clk,eq} = e_{clk} + e_{el,clk} \quad (11)$$

Indirect emissions are calculated as:

$$e_{el,clk} = e_{el} . P_{el,clk} \quad (12)$$

where $P_{el,clk}$ is the specific power consumption, and e_{el} is the CO_2 emissions associated with each unit of electric power consumed.

5 Results and discussion

Base plant without capture		Oxyfuel cement plant	
Parameter	**Value**	**Parameter**	**Value**
Clinker Production (t_{clk}/hr)	96.60	Clinker Production (t_{clk}/hr)	102.57
Cement Production (t_{clk}/hr)	131.08	Cement Production (t_{cem}/hr)	139.17
CO2 Emissions (t_{CO_2}/hr)	60.01	CO2 produced (t_{CO_2}/hr)	82.96
CO2 Concentration in Flue Gas (%)	15.89	CO2 Captured (t_{CO_2}/hr)	82.04
Specific CO2 Emissions (t_{CO_2}/t_{clk})	0.62	Specific CO2 Emissions (t_{CO_2}/t_{clk})	0.01
Cement Plant Thermal Requirement (MW)	107.68	Capture Efficiency (%)	98.89
Specific Heat Input (MJ/kg_{clk})	4.01	Cement Plant Thermal Requirement (MW)	107.68
Cost of Clinker (€/t_{clk})	51.22	Specific Heat Input (MJ/kg_{clk})	3.78
Cost of cement (€/t_{cem})	37.75	Cost of Clinker (€/t_{clk})	80.15
Cost of CO_2 Avoided	-	Cost of cement (€/t_{cem})	59.07
		Cost of CO_2 Avoided	50.68
		Flue gas recirculation ratio (%)	70

5.1 Technical Results

Table 6 Results from the technical and financial models

Table 6 shows a noticeable increase in clinker production as the plant with oxyfuel produced 102.57 t_{clk}/hr while the base plant without capture produced 96.6 t_{clk}/hr. This increase in clinker production is expected as the increase in volume fraction of oxygen in the combustion gas has been shown to lead to an increase in the clinker production rate given that the raw meal and fuel feed are kept the same. Equation 13 (Wrampe et al., 1976) shows the relationship between the amount of oxygen (Φ_2, in Million standard cubic feet per day), the fuel fed in, *G,* and the amount of clinker produced (*CR* in *tons/hour)*. The equation shows that the amount of oxygen is directly proportional to the amount to clinker produced.

$$CR = 281G + 39.1\Phi_2 - 12.8 \quad (13)$$

The CO_2 produced in the base plant was $60t_{CO_2}$/hr while that produced in the oxyfuel plant was $82.96t_{CO_2}/hr$. Although the oxyfuel plant produces more CO_2 than the base plant, $82.04t_{CO_2}/hr$ is captured in the CPU leading to a capture efficiency of 98.89% of the oxyfuel plant. The liquified CO2 stream has a purity of 98.2%.

5.1.1 Flue Gas Recirculation

Using pure oxygen with no recirculation increased the temperature of the kiln from 1738K in the base plant to approximately 3500K. This would have damaged the kiln and all subsequent units in the plant. It also would have resulted in extremely large heat exchangers to cool the clinker which would be expensive to install and operate. Therefore, a part of the flue gas was recycled back into the kiln and calciner via the first cooler to help reduce the temperature increase. As shown in Figure 2, the optimum recirculation rate to maintain temperatures similar to the base plant is around 70%, This stream has a composition of 81%(mass) CO_2 and 8.85%(mass) O_2.

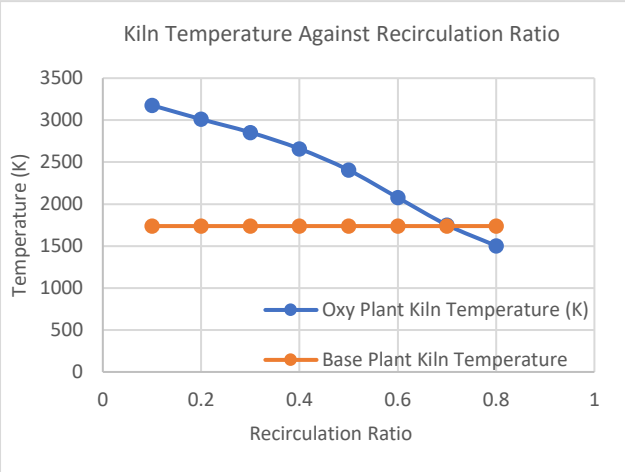


Figure 2 Kiln temperature varies with flue gas recirculation ratio

Hokfors et al. (2014) indicated that the ideal recirculation ratio was 60% but also noted that the oxyfuel plant's performance was similar to that of the plant without CCS when 80% of the flue gas was recycled with composition of approximately 80% wt. CO_2 and 10% wt. O_2. Granados et al. (2014)'s investigation of the combustion process with varying recirculation ratios and found that with a recirculation ratio of 72%, the flame length in the kiln of an oxyfuel plant was similar to that of a base plant without capture. While this model reached a conclusion similar to that of these two cases, the FGR ratio may vary with each plant, hence research before adoption is required.

5.1.2 Clinker Quality

Table 7 below shows the clinker compositions obtained from the base plant without capture and the oxyfuel cement plant. Dean et al. (2011) indicated that the typical ranges of the clinker phases are those shown in the table. As the mass compositions of clinker obtained from both models are similar and within the ranges provided by Dean et al. (2011), this suggests that the production of clinker by oxyfuel combustion technology probably will not affect the quality of clinker. This claim is also supported by Gerbelová et al. (2016) who stated that the overall quality of clinker is unchanged under oxyfuel conditions.

It is worth noting that tetracalcium alumino-ferrite is not present in both clinker compositions as it was not available in Aspen Plus databanks. This led to the significantly higher proportion of free lime (CaO) in both plants (4.3 wt% for plant without capture and 4.4 wt% for oxyfuel plant) as compared to the typical proportion of 1 wt%. The trace components in the models are also high due to the inability to include some of the reactions in the model. The presence of trace elements such as sulphur in actual clinker must be limited to quantities below 2.5%. Sulphur proportions greater than 2.5% lead to expansion and cracking of cement pastes upon hydration. An increase in sulphur has also been shown to cause a decrease in alite (responsible for long-term strength) and an increase in belite (responsible for short-term strength), which compromises cement strength. (Dean et al., 2010).

Table 7 Comparison of the clinker compositions from the base and oxyfuel plant

	Typical composition	**Base plant**	**Oxyfuel cement plant**
Component	**Mass fraction (%)**	**Mass fraction (%)**	**Mass fraction (%)**
Alite	50 to 70	61.7	61.3
Belite	15 to 30	15	15.3
Tricalcium aluminate	5 to 15	13.2	13.1
Tetracalcium alumino-ferrite	5 to 10	-	-
Free lime	1	4.3	4.4
Trace components	Varies	5.8	5.9

5.1.3 CPU performance

Table 8 Stream compositions showing CPU performance

CPU Performance		
Gas	**PPM at CPU inlet**	**PPM at cryo inlet**
H_2O	183000	316.00
NO	8.62	10.34
NO_2	0.0000196787	Trace
SO_2	2020.00	Trace
SO_3	1.04	Trace

The CO_2 purification unit is responsible for removing water, SOx and NOx from the CO_2 rich flue gas. The sour compression unit of the CPU was successful in removing SOx from the flue gas. SO_2 and SO_3 reduced from 2020ppm and 1.04ppm respectively to trace values leading to a 100% removal of the SOx (i.e. virtually none were left in the flue gas). This result is supported by Meunier et al. (2014) who showed that their SCU model resulted in 100% removal of SOx. The SCU, however, was not able to remove NO from the flue gas. This discrepancy will be studied further to investigate the ineffectiveness of the SCU to remove NO from the flue gas. Water was removed from the dehydration unit with an efficiency of 99.8% as the concentration in the flue gas decreased from 183000ppm to 316ppm.

5.1.4 Electricity usage

Table 9 Comparison of the electricity usage of the base plant and the oxyfuel cement plant

Reference plant without capture		Oxyfuel cement plant	
Component/Process	**Specific power consumption $P_{el,clk}$ (kWh/t_{clk})**	**Component/Process**	**Specific Power Consumption $P_{el,clk}$ (kWh/t_{clk})**
Raw meal and clinker grinding, solids handling, kiln drive, lightning etc	81.51	Raw meal and clinker grinding, solids handling, kiln drive, lightning etc	83.81
Coal preparation	4.64	Coal preparation	4.77
Fans		**Fans**	
ID fan	5.11	ID fan	3.75
Cooler fan	1.21	Cooler fan, blowers, recycle fans	12.31
Filter fan	2.42	Filter fan	1.78
Raw mill fan	10.13	Raw mill fan	10.11
Total fan electricity	**18.87**	**Total fan electricity**	**27.95**
Total electricity	**105.02**	ASU	37.44
		CPU	73.73
		Organic Rankine cycle	-38.22
		Total electricity	**189.47**

Table 9 show that the oxyfuel plant used 80.6% more electricity than the base plant without CO_2 (105.02 kWh/t_{clk} vs. 189.47 kWh/t_{clk}). The ASU and CPU were the main contributors to the significant increase in the consumption of electricity in the oxyfuel combustion plant. Overall, they each contributed 20% and 39% respectively to the total electricity consumption (Figure 3). The cooler, blower and recycle fans' electricity increased from 1.21 kWh/t_{clk} in the plant without capture to 12.31 kWh/$_{tclk}$ which also made a signification contribution to the increase in the oxyfuel plant's electricity consumption. On the other hand, the ID, filter and raw mill fans consumed marginally less electricity in the oxyfuel combustion plant (15.64 kWh/t_{clk}) than in the plant without capture (17.66 kWh/t_{clk}). This was owing to the reduced gas flowrate through each one of them as some of the flue gas stream was sent to the CPU.

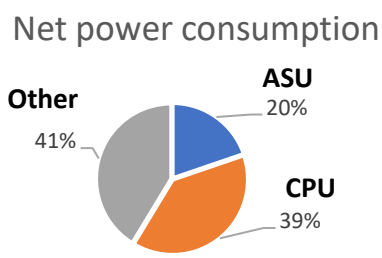


Figure 3 Net electricity usage of the ASU and CPU in the oxyfuel cement plant

The Organic Rankine Cycle produced 38.2 kWh/$_{tclk}$ of power from the waste heat. This was 16.8% of the total power which helped reduce the electricity costs. Electricity increase is one drawback of the oxyfuel technology.

Barker et al. (2009) showed an increment from 10MW in a plant without capture to 22MW to a partial oxyfuel combustion plant. IEA, 2013 showed a 128% increase when using a full oxyfuel combustion paint. This shows that the increment of power usage by 84.5kWh/t_{clk} is justified and further highlights the need for

comprehensive energy recovery in an oxyfuel combustion plant to help reduce long term costs on electricity.

5.2 Economic results

Table 10 Financial model results

	Base plant	Oxyfuel combustion
Variable OPEX M€/yr	15	35
Fixed OPEX M€/yr	18	21
Total capital requirement M€	207	323
Cost of clinker €/t_{clk}	51	80
Cost of CO2 avoided €/tCO_2	-	51

The total capital requirement of the base plant without capture was estimated to be 207 M€ while that of the oxyfuel combustion plant was estimated to be 323 M€ as shown by Table 10 on the left. This represents an increase in capital expenditure (CAPEX) of 56%. The increase in CAPEX is mainly due to the addition of the ASU and CPU, which account for 66.28 M€ of the 116 M€ increase in CAPEX. The ORC also accounts for 10.71 M€ of the increase in CAPEX as the base plant was not fitted with an organic Rankine cycle. Collectively the ASU, CPU and ORC account for 66.4% of the increase in CAPEX. The remaining 33.6% of the increase in CAPEX can be attributed to the recycle flue gas piping, modifications to the kiln, flue gas cooler and other engineering costs. The capital expenditure of 207 M€ of the base plant without capture compares closely to values suggested by CEMCAP (204 M€) and IEA (209 M€) of a similar plant configuration. The CAPEX of the oxyfuel plant of 323 M€ is also within a comparable range of values suggested by IEA (309 M€) and CEMCAP (323 M€) (CEMCAP, 2018) (IEA, 2013). These are shown in Figure 5 and Figure 6 in the supplementary material.
The fixed operating expenses increased from €15M a year in a plant without capture to €21M a year in the oxyfuel plant. This was mainly due to the increase, from 100 to 120, in the manpower. More labour was required for the operation and maintenance of the ASU and CPU. The units require specialised labour as they are relatively more technical and complicated.
Table 10 shows that the variable OPEX also increased by approximately 133% from the base plant to the oxyfuel plant. The increase was mainly due to the extra electricity used to operate the ASU and CPU as €8M per year (40% of the increase) is required for them. The other major contributor to the increase is the transport and storage of the CO_2 which accounted for approximately 25% of the €20M increase in variable OPEX. The rest of the increase is due to other new expenses like the cost of purchasing process water for the ORC and the TEG used for dehydration of the flue gas.

CEMPCAP (2018) and IEA (2013) models showed variable OPEX of €24M/yr and €28M/year respectively for the base plant without capture. These values are higher than the €15M/yr show in Table 10, this difference can be attributed to the inefficiencies of the bottom-up. The factor used in the six-tenths rule is an average value for different components thus there is a possibility of underestimating the cost of key components.

5.2.1 Cost of clinker

The cost of clinker for the base plant €51/t_{clk} while that of the oxyfuel cement plant is 80€/t_{clk}. The increase in the cost of clinker is due to the increase in annualized capital expenditure and the increase in O&M costs. IEA (2013) and CEMCAP (2018) suggested values of €74/t_{clk} and €63/t_{clk} for a similar base plant without capture and values of €104/t_{clk} and €93/t_{clk} respectively for a similar oxyfuel cement plant.

5.2.2 Cost of Avoided CO_2

The cost of avoided CO_2 (CAC) was found to be €$50.68/t_{CO_2}$. This calculation uses a power emissions factor of $0.49\ t_{CO_2}/kWh$ (Farrell, 2016) for all the electricity purchased externally. This value is lower than the CAC from the CEMPCAP (€ $42.40/t_{CO_2}$) and IEA (€$45.71/t_{CO_2}$) models. IEA's sensitivity analysis shows that, given the current prices and policies, the cost can be as high as €$52/\ t_{CO_2}$ therefore the value calculated from this model is within range. Given the underestimation of the cost of clinker for both the base and oxyfuel plant as shown in Table 10, this shows that there is a possible discrepancy in the calculation of the indirect emissions. This highlights the need for revalidation of the assumption made with regards to the emissions from the power acquired from the national grid and the inclusion of emissions from transport of cement and CO_2.

5.3 Sensitivity Analysis

Sensitivity analysis show that the variable operating costs have the most effect on the cost of clinker and CAC. A 25% variation in the variable OPEX resulted in a variation of €10.16/t_{clk} and €$17.80/t_{CO_2}$ in the CAC. Oxyfuel

cement plants can benefit from proximity to oxyfuel or coal gasification plants due to economies of scale in production of O_2 and treatment of flue gases.
The variation of fixed OPEX also caused a significant change to the CAC and the cost of clinker. Process automation can help reduce the number of employees on the plant and help reduce costs on labour. The accuracy of the power emissions factor is not a significant aspect as its variation by 25% only caused 5% variation in the CAC.

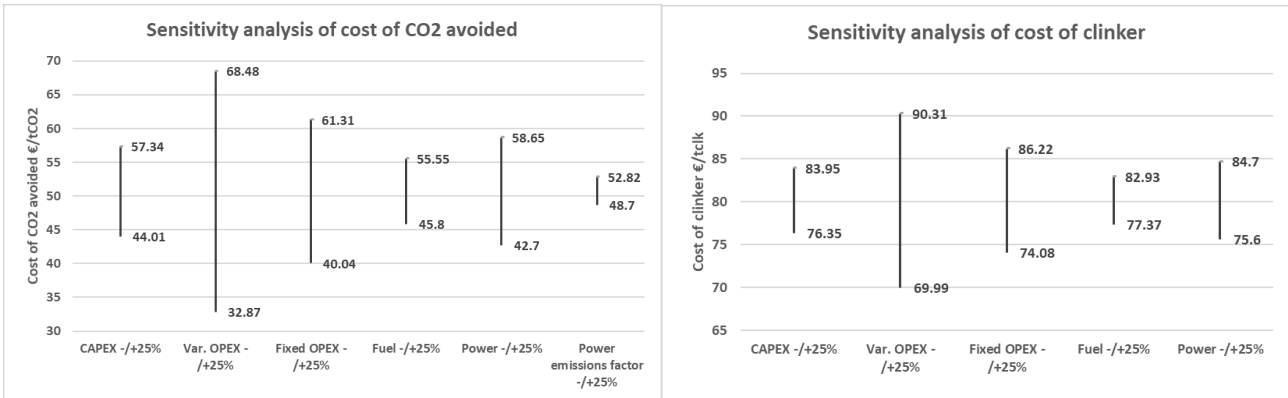


Figure 4 Sensitivity analyses of the cost of clinker and the cost of CO_2 avoided

6 Conclusion

This paper provides a comprehensive assessment of oxyfuel technology applied to a cement plant as a means for CO_2 abatement. The technical analysis was based on a model of a base plant without capture and a model of an oxyfuel combustion plant, both developed on Aspen Plus v9. The financial analysis was based on detailed process simulations and existing models.
The technical analyses of these plants show that the oxyfuel combustion plant produces more CO_2 ($82.96\ t_{CO_2}/hr$) compared to the base plant without capture ($60.01\ t_{CO_2}/hr$). The oxyfuel plant however was shown to have a CO_2 capture efficiency of 98.9% with a 98.2% purity of the CO_2 rich liquid stream. The use of oxyfuel technology in a cement plant does not influence the quality of clinker as the composition of clinker produced in both plant models was relatively the same. The optimum FGR ratio for the oxyfuel plant is shown to be 70% as supported by Granados e al. (2014) who showed that at a recirculation ratio of 72%, the flame length is similar to that of an air operated kiln.
The cost analysis indicates that the cost of clinker for the base plant without capture is €51.2/t_{clk} whereas that of the oxyfuel plant is €80.2/t_{clk}. These values are also shown, by the sensitivity analysis, to be mostly affected by the variable operating costs. The cost of CO_2 avoided for the oxyfuel plant is estimated to be €50.68/tCO_2. Although this value is slightly high than those suggested by IEA ($€46/t_{CO_2}$) and CEMCAP ($€42/t_{CO_2}$), it is still within a range of comparable values.
Moving forward, the models can be improved by including the models of the ASU, the raw meal treatment. Air ingress into the preheater, calciner and kiln should also be included to investigate the effects of air ingress on the clinker composition and kiln performance. Furthermore, the thermal energy requirement should be appropriately adjusted for each of the clinker manufacturing processes.

References

Ahmad, R., Khan, T.A., Agarwal, V. (2013) Mass and Energy Balance in Grate Cooler of Cement Plant. *International Journal of Scientific Engineering and Technology*. 2 (7), 631-637. Available from https://pdfs.semanticscholar.org/6e01/4c08a6d243ad9f61575031fa265788764160.pdf

Anantharaman, R., Berstad, D., De Lena, E., Fu, C., Gardarsdottir, S.O., Jamali, A., Pérez-Calvo, J., Romano, M., Roussanaly, S., Ruppert, J., Stallmann, O., Sutter, D., & Voldsund, M. (2018) CEMCAP comparative techno-economic analysis of CO2 capture in cement plants. *CEMCAP*. Available from http://doi:10.5281/zenodo.2597091.

Aspen Technology. (2013) Getting Started Modelling Processes with Solids. Available from http://profsite.um.ac.ir/~fanaei/_private/Solids%208_4.pdf

Banaszkiewicz, T., Chorowski, M., and Gizicki W. (2015) Comparative analysis of cryogenic and PTSA technologies for systems of oxygen production. *AIP Conference Proceedings*. 1573, 1373-1378. Available from https://doi.org/10.1063/1.4860866

Barker, D.J., Turner, SA., Napier-Moore, P.A., Clark, M. & Davison, J.E. (2009) CO2 Capture in the Cement Industry. *Energy Procedia*. 1 (1), 87-94. Available from https://doi.org/10.1016/j.egypro.2009.01.014.

Carrasco-Maldonado, F., Spörl, R., Fleiger, K., Hoenig, V., Maier, J., & Scheffknecht, G. (2016) Oxy-fuel combustion technology for cement production – State of the art research and technology development. *International Journal of Greenhouse Gas Control*. 45, 189-199. Available from https://doi.org/10.1016/j.ijggc.2015.12.014.

CEMBUREAU. (2017) Activity Report. Available from https://cembureau.eu/media/1716/activity-report-2017.pdf.
Chatterjee, A. K. (2018) Cement Production Technology Principles and Practice. 1st Edition ed.; Boca Ration.

Chemical Engineering Research and Design. 123, 230-239. Available from https://doi.org/10.1016/j.cherd.2017.05.013.

Cormos, A.M., & Cormos, C.C. (2017) Reducing the carbon footprint of cement industry by post-combustion CO2 capture: Techno-economic and environmental assessment of a CCS project in Romania.

Dean, C.C., Blamey, J., Florin, N.H., Al-Jeboori, M.J., & Fennell, P.S. (2011) The calcium looping cycle for CO2 capture from power generation, cement manufacture and hydrogen production. *Chemical Engineering Research and Design*. 89 (6), 836-855. Available from https://doi.org/10.1016/j.cherd.2010.10.013.

Dean, C.C., Dugwell, D., & Fennell, P.S. (2011) Investigation into potential synergy between power generation, cement manufacture and CO2 abatement using the calcium looping cycle. Energy & Environmental Science. (6), 2050-2053. Available from https://pubs.rsc.org/en/content/articlelanding/2011/EE/c1ee01282g#!divAbstract.

European Commission. (2001) Reference Document on Best Available Techniques in the Cement and Lime Manufacturing Industries.

Farrell, J.N. (2018) The Role of Industrial Carbon Capture and Storage in Emissions Mitigation. Massachusetts Institute of Technology. Available from https://globalchange.mit.edu/sites/default/files/Farrell_MS_2018.pdf

Gardarsdottire, S., De Lena, E., Romano, M., Roussanaly, S., Voldsund, M., Pérez-Calvo, J., Berstad, D., Fu, C., Anantharaman, R., Sutter, D., Gazzani, M., Mazzotti, M., & Cinti, G. (2019). Comparison of Technologies for CO2 Capture from Cement Production—Part 2: Cost Analysis. *Energies*. 12 (3). 542. Available from http://doi.org/10.3390/en12030542.

Gerbelová, H., van der Spek, M., Schakel, W. (2017) Feasibility Assessment of CO2 Capture Retrofitted to an Existing Cement Plant: Post-combustion vs. Oxy-fuel Combustion Technology. *Energy Procedia*.114, 6141-6149. Available from https://doi.org/10.1016/j.egypro.2017.03.1751.

Granados, D.A., Chejne, F., Mejía, J.M., Gómez, C.A., Berrío, A., & Jurado, W.J. (2014) Effect of flue gas recirculation during oxy-fuel combustion in a rotary cement kiln. *Energy*. 64, 615-625. Available from https://doi.org/10.1016/j.energy.2013.09.045.

Guo, Z. (2019) A Comparative Study between the Application of Calcium Looping Technology and the Use of Amine Scrubbing Technology to a Cement Plant.

Hökfors, B., Viggh, E., & Eriksson, M. (2014) Simulation of oxy-fuel combustion in cement clinker manufacturing. *Advances in Cement Research*. 27 (1), 42-49. Available from http://dx.doi.org/10.1680/adcr.13.00068.

IEA. (2009) Cement Technology Roadmap, Carbon emissions reductions up to 2050. Available from https://cement.mineralproducts.org/documents/wbcsd-iea%20cement%20roadmap%202009.pdf

IEAGHG. (2013) Deployment of CCS in the Cement Industry. Available from https://ieaghg.org/docs/General_Docs/Reports/2013-19.pdf.

IMARC Group. (2019) Cement Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2019-2024. Available from https://www.researchandmarkets.com/research/gcjn2b/global_cement?w=4

Laribi, S., Dubois, L., De Weireld, G., & Thomas, D. (2017) Optimization of the Sour Compression Unit (SCU) process for CO_2 Purification Applied to Flue Gases Coming from Oxy-combustion Cement Industries. *Energy Procedia*. 114, 458-470. Available from https://doi.org/10.1016/j.egypro.2017.03.1188.

Łukowicz, H., & Kochaniewicz, A. Analysis of the use of waste heat obtained from coal-fired units in Organic Rankine Cycles and for brown coal drying. Energy. 45 (1), 203-212. Available from https://doi.org/10.1016/j.energy.2012.03.035.

Meunier, N., Laribi, S., Dubois, L., Thomas, D., & De Weireld, G. (2014) CO_2 Capture in Cement Production and Re-use: First Step for the Optimization of the Overall Process. *Energy Procedia*. 63, 6492-6503. Available from https://doi.org/10.1016/j.egypro.2014.11.685.

Nivargi, J.P., Gupta, D.F., Shaikh, S.J., & Shah K.T. (2005) TEG Contactor for Gas Dehydration. Chemical engineering world. 40 (9), 77-80. Available from http://www.fenixchemtech.in/pdf/teg_contactor.pdf.

Rahman, A., Rasul, M.G., Khan, M.M.K., & Sharma S.C. (2017) Assessment of Energy Performance and Emission Control Using Alternative Fuels in Cement Industry through a Process Model. *Energies*. 10 (12), 1996. Available from https://doi.org/10.3390/en10121996.

Rahman, A., Rasul, M.G., Khan, M.M.K., & Sharma, S. (2014) Aspen Plus Based Simulation for Energy Recovery from Waste to Utilize in Cement Plant Preheater Tower. *Energy Procedia*. 61, 922-927. Available from https://doi.org/10.1016/j.egypro.2014.11.996.

Rodgers, L. (2018) *Climate change: The massive CO2 emitter you may not know about*. BBC.

Schorcht, F., Kourti, I., Scalet, B.M., Roudier, S., & Sancho, L.D. (2013) Best Available Techniques (BAT) Reference Document for the Production of Cement, Lime and Magnesium Oxide. *European Commission*. Available from http://doi.org/10.2788/12850.

Adsorption Studies in Nanoporous Materials Using the Zero Length Column Technique

Matthew Farmer
matthew.farmer16@imperial.ac.uk

Andrea Rovelli
andrea.rovelli16@imperial.ac.uk

Abstract

The zero length column (ZLC) is a simple and flexible technique allowing for rapid characterisation of adsorbents. Although its versatility has been proven over the past decade, the technique has never been attempted within Imperial College London. This paper aims to address both the ZLC's theory and the development of an experimental prototype for equilibrium analysis. A mathematical model allowing for the implementation of the ZLC theory was developed within Matlab; flexibility of which was shown in the simulation of virtual experiments. A prototype setup was built and experiments of CO_2 adsorption on Norit RB3 Activated Carbon successfully displayed operation at equilibrium conditions and experiment repeatability. A Henry constant for the adsorbent was calculated and found to suitably agree with that predicted by a virtual experiment. A systematic error and associated limitations were also identified along with recommendations for future experimental setups.

I. Introduction

The design of adsorption and catalytic reaction processes requires knowledge of the transport properties of the adsorbed species on the microporous solid involved (Brandani et al., 1996). On the other hand, processes such as pressure swing, require equilibrium adsorption properties. Methods of obtaining both these properties exist, such as the Wicke-Kallenbach method and NMR methods for kinetic parameters (Ruthven, 1984), and volumetric and gravimetric methods for equilibrium parameters (Keller et al., 2005). However, due to the time consuming nature of these methods, and the specificity of the experimental setup not offering large flexibility, a method allowing for the determination of both equilibrium and kinetic parameters is highly attractive. Originally developed in the work by Mladen et al. (1988), the zero length column (ZLC) method provides a simple, rapid, and relatively inexpensive method for obtaining these physical properties (Hu et al., 2015a).

The ZLC method involves equilibration of a small sample of adsorbent with a known sorbate gas mixture, the sample is subsequently flushed with a pure purge gas stream and the desorption curve is then measured (Krishnamurthy et al., 2019). By altering the experimental conditions, commonly via the purge gas flowrate, the system can be operated under either equilibrium or kinetic control - enabling the calculation of both equilibrium and kinetic parameters (Friedrich et al., 2015).

Due to the increase in interest in topics such as carbon capture and storage, techniques for rapid screening of adsorbents are increasingly sought after. The ability to quickly assess the adsorption properties, from both an equilibrium and kinetic standpoint, between CO_2 and a newly developed adsorbent would allow for higher throughput of new adsorbents and, as such, faster developments in carbon capture technology (Hu et al., 2015b). The flexibility in the screening is additionally complemented by the ZLC's ability to function using very small samples - making it ideal for testing newly designed adsorbents.

i. Background

The ZLC technique was initially designed to measure the diffusivity of pure gases in microporous solids, since then it has been effectively utilised within a vast number of different research areas; with extensive work done in the measurement of diffusivity in zeolites (Hu et al., 2015b). The kinetics of processes controlled by intra-crystalline diffusion (Grenier et al., 1994), macropore diffusion (Hu et al., 2014) and surface resistance (Vidoni et al., 2012) have all been explored with the ZLC method. Multiple variants of the technique have been developed for specific measurements, these include the tracer ZLC (Brandani et al., 1995), the counter-current ZLC (Brandani et al., 2000), and the partial loading ZLC experiment (Gunadi et al., 2006). The theory supporting the technique has also been thoroughly investigated, with models developed for application to nonlinear systems (Brandani, 1998a), non-isothermal systems (Brandani, 1998b), biporous systems (Brandani, 1996), and systems with crystal size distribution (Duncan et al., 2002). Furthermore, successful measurements of systems under equilibrium control were performed, allowing equilibrium isotherms to be extracted directly from desorption curves (Brandani et al., 2003). This particular application has established the ZLC as a vital tool in investigating adsorption properties of newly developed adsorbents (Mangano et al., 2013a).

ii. Theory

The vast majority of the theory developed regarding the ZLC technique revolves around the mass balance

within the column, given by Equation (1), where, due to the small size of the column, a perfectly mixed cell assumption is made (Brandani et al., 2003).

$$V_s \frac{dq}{dt} + V_g \frac{dc}{dt} + Fc = 0 \tag{1}$$

In Equation (1), the three terms can be attributed to individual contributions within the ZLC column. The first term, comprised of V_s, the volume of the sample adsorbent, and the time derivative of q, the adsorbed phase concentration, is the term attributable to the adsorbent itself. The second term, containing V_g, the dead volume of the experimental setup, and the time derivative of c, the concentration of sorbate in gas phase, represents the contribution to the mass balance from the gas volume surrounding the adsorbent. The last term is the convective contribution, consisting of F, the flowrate of purge gas, and c. Brandani et al. (1995) showed that an analytical solution can be given by Equation (2), where fluid-phase holdup is neglected, Fickian diffusion of diffusivity D and linear equilibrium isotherm with Henry constant K are assumed.

$$\frac{c}{c_0} = 2L \sum^{\infty} \frac{\exp -\beta_n^2 (\frac{D}{r^2})t}{\beta_n^2 + L(L-1)} \tag{2}$$

Here β_n is given given by the positive root of Equation (3) and L by Equation (4).

$$\beta_n \cot \beta_n + L - 1 \tag{3}$$

$$L = \frac{1}{3} \frac{F}{KV_s} \frac{r^2}{D} \tag{4}$$

The dimensionless parameter L can be best interpreted by viewing its two contributions: convective desorption time constant, (F/KV_s), and diffusional time constant, (D/r^2) (Friedrich et al., 2015). As L represents the ratio between the two time constants, it follows that, for small values of L, one can assume that equilibrium between purge gas and adsorbent is achieved, resulting in convection controlled desorption. On the other hand, when L is large, kinetic controlled adsorption is present.

iii. Motivation and Objective

This research project aims to achieve two objectives. Firstly, it seeks to explore the current theory surrounding the ZLC technique and demonstrate the utility in understanding this theory. Secondly, it seeks to develop a prototype ZLC system to allow for the experimentation of the adsorption of CO_2 on pelletised activated carbon. In achieving these two objectives, a foundation for the continuation and progression of this experimental technique within *Imperial College London* is then built. This is particularly attractive due to the aforementioned flexibility the ZLC technique offers; allowing for, *inter alia*, rapid ranking of adsorbent capabilities (Hu et al., 2015b), and measurements of diffusive properties of adsorbents (Gelles et al., 2020). Acknowledging that the past four years have contributed to approximately 45 % of all research published regarding 'new adsorbent' (Clarivate Analytics, 2019), it is clear to see the appeal and utility of having a technique which compliments this rise in popularity. To thus achieve the ZLC's full potential, its fundamentals must first be explored and understood.

II. Numerical Investigation Methods

One major aspect in building a fundamental understanding of the ZLC technique is to familiarise oneself with the theory and develop an appropriate mathematical model that allows for both computational exploration and experimental analysis.

As was outlined in Section I.ii, the ZLC theory can be analysed under two regimes, equilibrium regime and kinetic regime. For the purpose of this research, equilibrium regime was investigated - signifying $L < 1$. Utilising this constraint on L, yielding $\beta_n^2 \to 3L$ from Equation (3), and assuming longer times, such that only the first term of Equation (2) is considered, one can show that Equation (2) reduces to Equation (5) (Mladen et al., 1988).

$$\frac{c}{c_0} = \exp \frac{-Ft}{KV_s + V_g} \tag{5}$$

Here the definition of L from Equation (4) was used and V_g was introduced to account for the dead volume of the column. Interestingly, Equation (5), can more easily be derived from the integration of Equation (1), with $\frac{dq}{dt} = \frac{dq}{dc}\frac{dc}{dt}$ and $q = Kc$, equivalent assumption to that made in the derivation of Equation (2).

Equation (5) indicates that, under equilibrium conditions, c/c_0 is invariant of Ft. This allows one to determine whether their experimental results are under equilibrium condition by plotting c/c_0 vs Ft for differing purge gas flowrates and identifying whether the response curves collapse - signifying invariance.

By definition, under the equilibrium assumption, the sorbate concentration in the gas and the sorbate concentration in the adsorbent are equal. It is then possible to recover the isotherm from the Ft response curves using Equation (6) (Brandani et al., 2003).

$$q^* = \frac{1}{V_s} \int_t^{\infty} \frac{c/c_0}{1/y_0 - c/c_0} dFt - \frac{V_g}{V_s} c \tag{6}$$

Equation (6) has been shown to be valid up to a y_0, initial sorbate mole fraction, of 0.1; after which, the assumption of $F_{\text{in}} = F_{\text{out}}$ is no longer valid due to the contribution to the exiting flowrate from the desorbing sorbate (Mangano et al., 2013b). The integral term in Equation (6) is the total amount of sorbate within the adsorbent at time t, whilst the second term accounts

for the sorbate present in the dead volume of the column.

i. Mathematical model

Having outlined the required background relating to the theory, it is clear to see that for ZLC experiments under equilibrium conditions, there are two main domains of knowledge: isotherms and c/c_0 vs Ft plots. The mathematical model section will thus aim to create bridges between the two domains. The link relating isotherms to c/c_0 vs Ft plots will be achieved via Equation (1), where, setting appropriate experimental conditions, namely $[V_s, V_g, F, c_0, q]$, $[c, t]$ will be calculated via numerical integration using ode45, based on the Runge-Kutta method, within *Matlab*. On the other hand, the bridge between c/c_0 vs Ft plots to isotherms will be built using Equation (6). By defining $[c, t]$ and the corresponding conditions, namely $[V_s, V_g, F]$, the isotherm q can be obtained via numerical integration using trapz, based on the trapesodial method, within *Matlab*.

ii. Experimental analysis

Once suitable links between the two domains of knowledge have been created, experimental analysis can easily be carried out. Experimental results take the form $[c, t]$, naturally lending themselves to be represented via c/c_0 vs Ft plots. Using these plots, Equation (5) can be utilised to derive the Henry constant for the adsorbent, K. For this to be the case, V_g must first be determined using blank experimental runs, where no adsorbent is present in the column ($V_s = 0$) - resulting in Equation (5) being reduced to a singular unknown. Equation (5) calls for the exponential fitting of the experimental data; this is achieved via Equation (7).

$$\min_{\mathbf{x}} \sum_{i=1}^{N} \sum_{j=1}^{M} \left(\frac{S_{i,j} - P_{i,j}(\mathbf{x})}{S_{i,j}} \right)^2 \tag{7}$$

Here N is the number of experimental runs being simultaneously fitted, M is the number of data points within each run, $S_{i,j}$ is the experimental data point corresponding to run i and point j, and $P_{i,j}$ is the predicted value for point $S_{i,j}$ and fitting parameters $\mathbf{x}$. For the purpose of the optimisation, an in-house particle swarm optimiser built in *Matlab* was used (Haupt et al., 2003), where inertia and constriction factors were set via recommendation from the work of Eberhart et al. (2000). This fit then allows for calculation of K, whilst the mathematical model formerly described allows for calculation of the isotherm. Due to the application of simultaneous fitting, uncertainties for the Henry constant were then calculated as the difference between the simultaneously fitted calculated value and the average of the calculated values from individual fits.

III. Experimental

Owing to the simplicity of the technique, the ZLC experimental setup is equally minimal. The equipment can be categorised into three distinct segments: initial gas feed, ZLC cell, and gas detection/identification. Within the experimental investigation, two iterations of the setup were tested, shown by the P&IDs in Figure 1; however, meaningful results were only extracted using setup **B** due to the smaller dead volume - all focus will thus be placed on this setup.

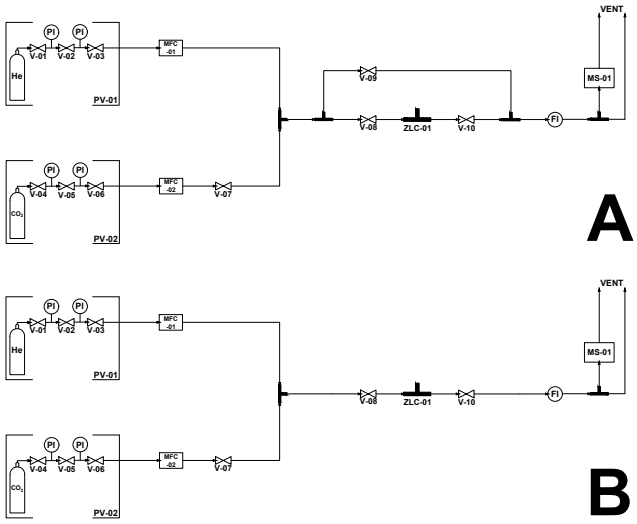


Figure 1: P&ID of the two setup iterations tested within the experimental research. Setup **B** was found to provide more meaningful results due to the smaller dead volume.

The gas feed section consists of two gas canisters, one containing pure He and the other pure CO_2. The gas flow from each canister was controlled by individual mass flow controllers (*Alicat Scientific Mass Flow Controllers*), both gas streams then connected to a 1/8" Swagelok T-piece. The ZLC cell itself was a 1/4" Swagelok T-piece with capped top and, to ensure that the activated carbon pellet remained secure within the cell, a metal gauze fitted at the point of entry and exit of the purge gas. An accurate flow indicator (*Alicat Scientific Mass Flow Meter*) was positioned before the mass spectrometer to both calibrate the mass flow controllers, and to check for leaks throughout experimental runs. The mass spectrometer used (*Pfeiffer Vacuum OmniStar*) provided live gas composition readings and recorded the desorption curves from the experiments conducted.

The adsorbent used for the experiments was *Norit RB3 Activated Carbon*, a pelletised activated carbon adsorbent. The pellet was found to have a mass of 0.0264 g, which, via the skeletal density, corresponded to a volume of sample, V_s, of 0.0136 mL.

Before commencing the first experimental runs, the two mass flow controllers had to be calibrated indi-

vidually. For calibration of both the mass flow controllers and mass spectrometer, the ZLC cell along with surrounding valves (V-08, ZLC-01, and V-10) were removed and set in an oven at 120 °C for adsorbent regeneration. Pure He was then ran through MFC-01 at incremental flowrates, the corresponding flowrate would be noted from the flow indicator, resulting in a calibration curve. The same process was followed for pure CO_2 and MFC-02. With mass flow controllers calibrated, the mass spectrometer could then be calibrated daily using an equimolar stream of He and CO_2.

Once calibration was completed, the ZLC cell is then reinstated and the system is equilibrated with a known He/CO_2 gas mixture of 6 to 8 vol% CO_2. Once the CO_2 concentration reading was stable on the mass spectrometer, indicating saturation of adsorbent and equilibration, the CO_2 supply was shut off, valve V-07 was closed, and the He flowrate was set to the specified purge flowrate. The desorption curve was then recorded on the mass spectrometer until a CO_2 concentration of below 0.1 vol% was achieved. The system was then re-equilibrated with a known CO_2 concentration and the process was repeated for multiple purge flowrates. Due to the nature of our setup, and aim to develop a simple first prototype, all experimental runs were performed under room temperature.

IV. Results

i. Mathematical model

The numerical investigation aimed to create bridges between the two domains of knowledge within the ZLC technique: isotherms and $\frac{c}{c_0}$ vs Ft plots. To maintain generality and demonstrate maximum flexibility, three different isotherm types were selected for this investigation, shown in Figure 2 as solid lines. A Langmuir isotherm (monolayer isotherm), a Henry isotherm (simple linear isotherm), and a BET isotherm (multilayer isotherm), were chosen due to their evident differences. Equation (1) was then applied in order to simulate an experimental run for all three isotherms, this yielded results of the form $[c, t]$ - shown in Figure 3 as $\frac{c}{c_0}$ vs Ft plots. Having obtained $\frac{c}{c_0}$ vs Ft data, Equation (6) was then used to calculate the corresponding isotherms - shown in Figure 2 as markers.

Notably, Figure 3 shows that the shape of the response curves is dependent on the isotherm used in the data generation - signifying the ability of $\frac{c}{c_0}$ vs Ft plots to retain information regarding isotherm type. This is particularly important as, for experiments with new adsorbents, knowledge of the isotherm type will, most likely, not be known *a priori*; having the ability to determine the isotherm type directly from the response curve is thus appealing.

It is clear from the results, particularly from acknowledging the perfect overlap of original isotherms and calculated isotherm points in Figure 2, that the mathematical links between the two domains of knowledge was successful. This is pivotal in having confidence in ones experimental results as, by having a proven system of moving between isotherms and $\frac{c}{c_0}$ vs Ft plots, uncertainty over correct interpretation of data disappears - one then has the guarantee that the

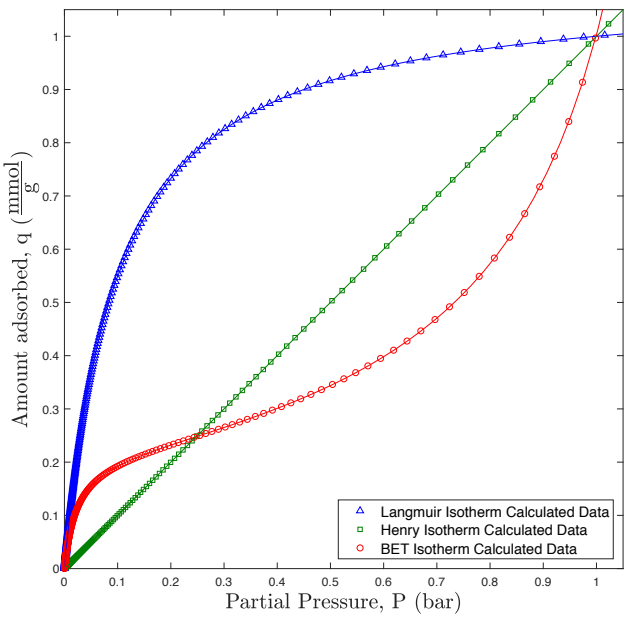


Figure 2: Three types of isotherms used within the numerical investigation. Original isotherms used in data generation shown in solid line and calculated isotherm points shown with markers.

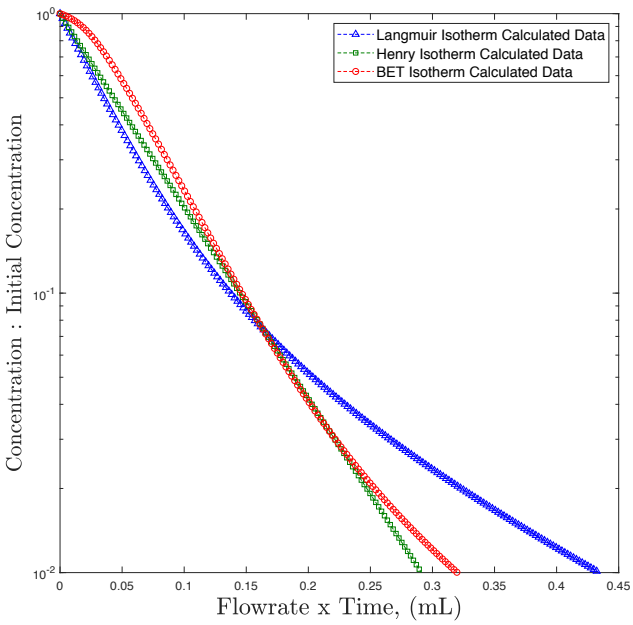


Figure 3: $\frac{c}{c_0}$ vs Ft plot for the three types of isotherms from numerical investigation. Data generated using isotherms from Figure 2 in combination with Equation (1).

results presented are correctly calculated.

ii. Experimental

Experimental analysis requires system characterisation via the dead volume of the experimental setup. Blank runs were undertaken and the results are shown in Figure 4.

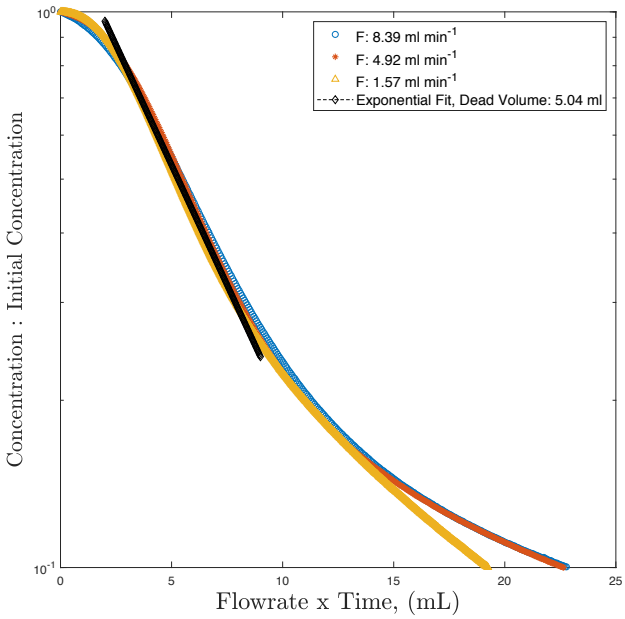


Figure 4: Blank runs: experimental results, exponential fit and calculated dead volume.

The exponential fit reported is in accordance with Equation (5) with $V_s = 0$, as per blank run definition. The calculated dead volume of the system, V_g, was thus found to be 5.04 mL.

Figures 6 reports the experimental results as a $\frac{c}{c_0}$ vs t plot for the activated carbon pellet being investigated. Figure 7 displays this data in a $\frac{c}{c_0}$ vs Ft plot - allowing for more meaningful observations. By fitting Equation (5) to the $\frac{c}{c_0}$ vs Ft plots it is then possible to extract a Henry constant for the adsorbent; this was calculated as $(7.1 \pm 0.2)\,\mathrm{mmol\,g^{-1}\,bar^{-1}}$. Additionally, by virtue of Equation (6), Figure 5 reports the calculated isotherms for the different experimental runs.

V. Discussion

Having shown the validity of the mathematical method employed within the numerical investigation, it is possible to confidently, in relation to the numerical techniques employed, interpret the experiment results. The aim of the experimental aspects of the research project was to develop a prototype ZLC system for experimentation; this section thus aims to analyse the results in that regard.

i. System and Domain Characterisation

Firstly, from Figure 4, it is interesting to note the convergence of the $\frac{c}{c_0}$ vs Ft plots. Albeit no adsorbent

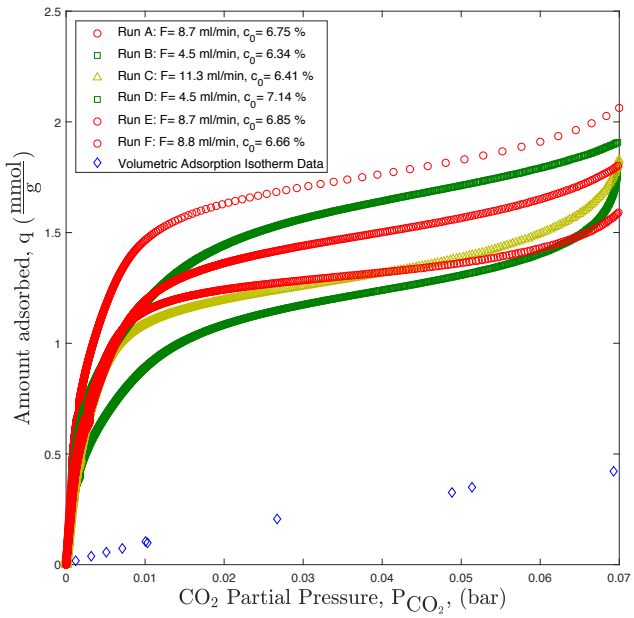


Figure 5: Calculated isotherms for experimental results. Also shown is isotherm data for the same adsorbent obtained via volumetric method.

is present in blank runs, one still expects the dead volume of the system to be independent of the flowrate of the purge gas stream; the convergence of the curves represents this and is an indication of a sound experimental results. The dead volume calculated for the system was 5.04 mL. As was alluded to in Figure 1, this value was larger, 5.95 mL, when setup **A** was initially used. Despite this reduction in dead volume, the dead volume for the experimental prototype remained considerably larger than those found in literature, ≈ 0.15 mL (Brandani et al., 2003); the effect of which is discussed later.

With the experimental system characterised by the dead volume, experimental runs with adsorbent present are reported in Figures 6 and 7. The combination of the two figures help illustrate the concept of invariance in Ft; showing a collapse of curves when plotted as $\frac{c}{c_0}$ vs Ft for experimental runs of differing purge flowrate, as Equation 5 implied. Although this result might appear trivial, it is undoubtedly the most important result within ZLC equilibrium studies. It was stated that the ZLC's experimental regime depended on the dimensionless parameter 'L', from Equation (4); however, specifically for new adsorbents, it is often impossible to calculate a value for 'L' *a priori*. For equilibrium studies, one thus commonly operates at low purge flowrates and subsequently verifies whether convergence of $\frac{c}{c_0}$ vs Ft plots occurs - ensuring oneself of operation in the intended regime. Should this be found to not be the case, subsequent analysis regarding isotherm derivation is no longer applicable, and any associated result would be meaningless. Another, less significant, albeit equally important, point

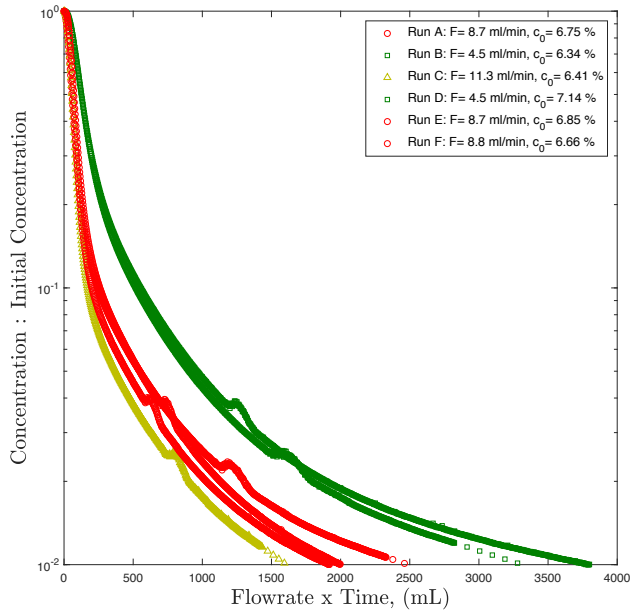


Figure 6: $\frac{c}{c_0}$ vs t plot for experimental results - higher purge flowrates intuitively showing a faster decay in concentration.

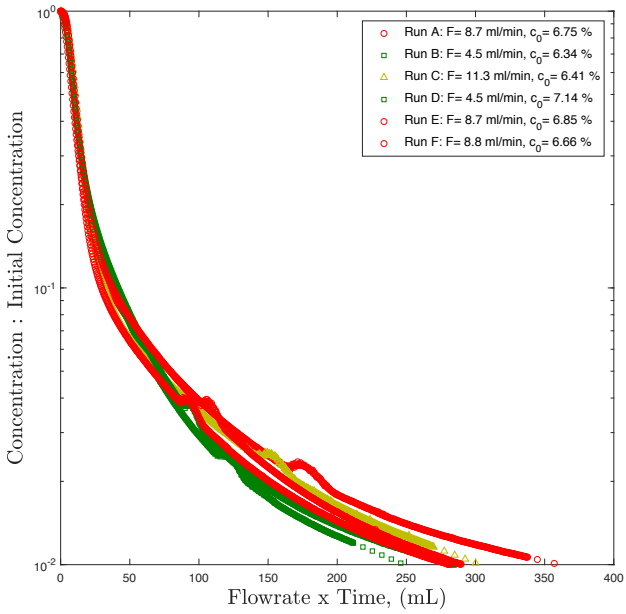


Figure 7: $\frac{c}{c_0}$ vs Ft plot for experimental results - convergence of curves is evident indicating equilibrium conditions.

of interest regards Figure 6. Here one can see that, for runs of similar initial and operational parameters, shown grouped by colour, the resulting $\frac{c}{c_0}$ vs t plots strongly overlap; acknowledging these results were obtained over different days and, by virtue, with differing mass spectrometer calibration factors, this implies repeatability of the results - an extremely important factor when assessing the performance of a prototype experimental setup.

ii. Identification of Systematic Error

Having determined that both the intended domain of operation, equilibrium control, was achieved, and that the results were repeatable, subsequent analysis regarding isotherms is justified and sound. Utilising Equation (6), in combination with the proven bridge developed in the mathematical model, Figure 5 reports the corresponding isotherms. Two observations can be rapidly made, firstly, there exists a poor convergence between the calculated isotherms and, secondly, there is a large overestimation of the amount of sorbate adsorbed. In relation to the first point, by reminding oneself that all experimental runs were carried out with the same adsorbent, it is no surprise that one would expect a single 'true' isotherm, inherent to the adsorbent, on which all calculated isotherms would collapse - here this is, evidently, not the case. In relation to the second point, Figure 5 also reports isotherm data for the same adsorbent obtained via volumetric methods, which was additionally cross-checked with the work of Kumar et al. (2018); it is clear that, via comparison, all calculated isotherms largely overestimate the amount adsorbed - implying the existence of a systematic error within the experimental prototype setup. As a bold statement, this conclusion can only confidently be made as a result of our knowledge that the mathematical tool used to calculate the isotherms is sound. As such, the error in the calculated isotherms, depending solely on the numerical approach taken and the experimental data itself, must be attributable to the experimental data - determining this error is thus pivotal in understanding the level of success of the prototype built.

iii. Determination of Systematic Error

Owing to the number of unknown influences, determining a source of error is often challenging; however, major strides can be made by employing the mathematical model built within the numerical investigation of this report and undertaking an important virtual experiment. By utilising the combination of experimental system characterisation and volumetric isotherm data, calculation of the expected response curve for the adsorbent is then possible. This is done as follows; firstly, the volumetric isotherm data is fitted to a Langmuir isotherm, shown as a solid black line in Figure 8. Secondly, the numerical bridge between isotherm and $\frac{c}{c_0}$ vs Ft plots from the mathematical model is used to generate an expected response curve, shown as blue markers in Figure 9. Lastly, for the sole purpose of validation, the numerical bridge between $\frac{c}{c_0}$ vs Ft plots and isotherms from the mathematical model is utilised to calculate the isotherm of the generated response curve, shown in Figure 8 as blue markers perfectly converging with the original Langmuir fit - indicating a successful

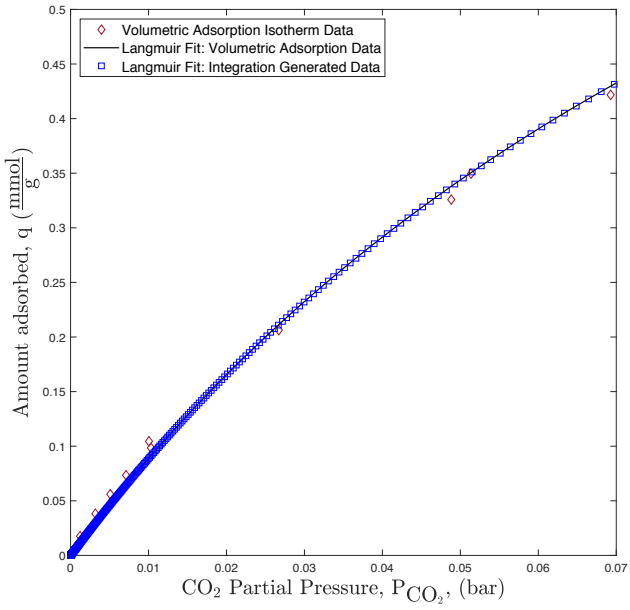


Figure 8: Isotherm showing original volumetric data points, Langmuir fit and corresponding calculated isotherm from generated response curve in Figure 9.

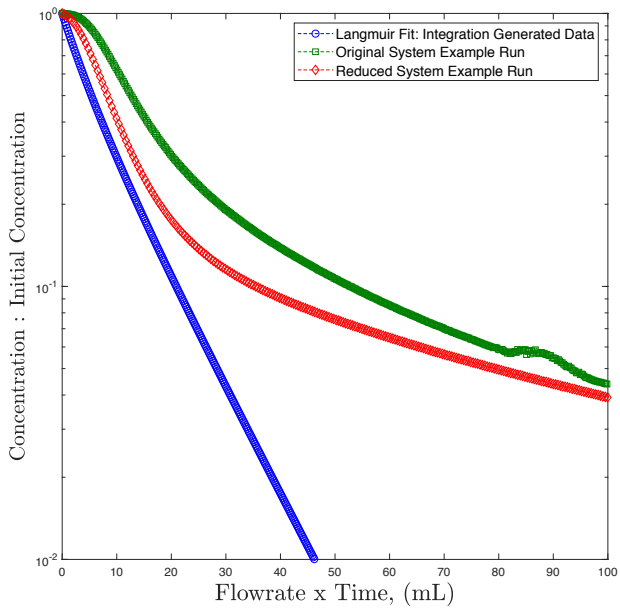


Figure 9: $\frac{c}{c_0}$ vs Ft plot for expected response curve of adsorbent investigated, calculated from volumetric adsorption data, and experimental results from this work.

virtual experiment.

Having obtained an expected response curve for the adsorbent investigated, and having confidence in the mathematical method employed, it is now possible to examine the results and attempt to determine the source of experimental error within the prototype built. From Figure 9 three observations are evident: firstly, the experimental data appears to display an asymptotic 'tail-like' behaviour in the longer time region, secondly, the experimental data appears to display concavity in the $\frac{c}{c_0}$ vs Ft plot for low Ft and, lastly, the experimental data appears to show greatest confidence in the initial slope of the curves. Analysis of the first two observations will yield information regarding the error within the experimental prototype whilst analysis of the last observation will display the capabilities of the system bounded by limits discovered.

The first two observations can be plausibly explained by assuming the system had both a large dead volume, already stated to be true when compared to literature, and a ZLC cell too large for the adsorbent sample. It was mentioned that, in the derivation of Equation (1), a well-mixed cell was assumed. This is because, for the isotherm to be calculated from the desorption curve, the concentration of the purge gas exiting the ZLC cell must be perfectly equal to that of the adsorbent; should any purge gas bypass occur, this assumption is no longer valid. By having a ZLC cell too large for the adsorbent, bypass becomes a possibility. This fact, coupled with the large dead volume of the system, allowing for stagnation, could result in areas of uneven concentrations; areas of high concentration due to stagnation, and areas of lower concentrations due to purge gas bypass. The resulting possibility of diffusion from the areas of stagnation, particularly evident in the long time region due to the overall lower concentration of the system, would be displayed as a more gradual concentration decay, or tail like behaviour. The initial concavity of the curves can also be similarly explained; the concavity signifies a delayed, in respect to the total amount of purge gas inputted in the system (Ft has units mL, it is the volume of purge gas passed through the system), decrease in concentration. This delay can be seen as the initial purge gas entering the system, bypassing the adsorbent, and exiting the system - another indicator of bypass. Overall, the most probable causes of the experimental error are thus both the possibility of purge gas bypass and the large dead volume of the system; as, combined, these two factors adequately describe the experimental observations.

The last observation, relating to the experimental results displaying an area of greatest confidence in the initial slope, allows for two interesting results. Firstly, it was mentioned how application of Equation (5) to the initial slope allows one to calculate the adsorbent's Henry constant; this is still possible - results shown in Table 1. Here results for the original system, setup **A**, and the volumetric isotherm data, via the generated response curve, are also shown.

The results in Table 1 show how, despite the aforementioned error in the experimental system affecting the long time region elements and extremely short time

Table 1: Calculated Henry constants for experimental runs and expected response curve of *Norit RB3 Activated Carbon.*

Henry Constants	$\mathrm{mmol\,g^{-1}\,bar^{-1}}$
Volumetric isotherm data	6.6
Original system	9.3 ± 0.3
Reduced volume system	7.1 ± 0.2

region elements, the initial slope remains a reliable source of data. Albeit not yielding the exact value for the Henry constant investigated, acknowledging that the ZLC technique has primary potential in the rapid screening of new adsorbents, the capability of achieving a good representative value for the Henry constant, via a set of experimental runs totalling less than one day, is a major achievement.

The last of these two results relates to the limits of the experimental setup built. This relates to the juxtaposition between Figures 3 and 9; in both figures a $\frac{c}{c_0}$ vs Ft plot for a Langmuir isotherm is shown, however, clear differences are evident. The response curve from Figure 3 displays evident curvature whilst that in Figure 9 appears linear; this is attributable to the differences in the dead volumes of the simulations. Figure 9 employed a dead volume equal to that of the experimental prototype, 5.04 mL, whilst Figure 3 used a substantially smaller dead volume, ≈ 0.05 mL. If one acknowledges that an analytical solution of Equation (1) for a Langmuir isotherm exists, given by Equation (8) (Brandani et al., 2003), and that Equation (8) can be subsequently fitted to the $\frac{c}{c_0}$ vs Ft plot - allowing for the extraction of both K and b - one can also acknowledge that, with a large dead volume causing loss of curvature in the response curve, this fitting is no longer possible.

$$\ln\left(\frac{c}{c_0}\right) = \frac{-Ft}{KV_s + V_g} - \frac{KV_s}{KV_s + V_g}\left[\frac{1}{1+bc} - \frac{1}{1+bc_0} + \ln\left(\frac{1+bc_0}{1+bc}\right)\right] \quad (8)$$

This implies that, for experimental systems where the dead volume is too large, extraction of certain isotherm parameters, namely all parameters bar Henry constants, becomes impossible - as is the case for our system.

VI. Conclusion and Future Work

This work presented two objectives: demonstrate the theory surrounding the ZLC technique and develop a prototype ZLC system for experimentation. The applicability of the theory was shown via development of a mathematical model which allows one to transition between the ZLC's two domains of knowledge, isotherms and $\frac{c}{c_0}$ vs Ft plots; this tool's flexibility was then demonstarted through the use of virtual experiments. A prototype ZLC system was also built for experimentation; this allowed for confirmation of the invariance in Ft for plots of $\frac{c}{c_0}$ vs Ft under equilibrium conditions and calculation of a Henry constant for the pelletised activated carbon. Experiments were also shown to be repeatable, indicating a sound experimental system and method. A systematic error was identified by analysing the calculated isotherms and probable cause attributed to the combination of both the large dead volume of the experimental system and a ZLC cell too large for the adsorbent.

i. Future Experimental Setups

One important aspect of the work presented in this paper is the learning aspect inherent to building a first prototype of an experimental setup. It was shown that the setup, considering its use in rapid screening of adsorbents, yielded suitable values for the Henry constant of pelletised activated carbon; however, it was also shown that the experimental system was limited in only allowing for the extraction of this parameter due to the large dead volume. Future experimental setups should thus aim to improve on this aspect by attempting to minimise dead volume of the experimental system and using a smaller ZLC cell to host the adsorbent investigated. This is a result of the visibility of the response curve's curvature depending on both the dead volume and the adsorbent's isotherm, when investigating novel adsorbents - in which the isotherm is unknown - to maximise the likelihood of curvature being visible in the response curves, the dead volume should thus be minimised. On the other hand, should the adsorbent's isotherm be known, one could employ the mathematical model developed to obtain an estimate for the dead volume necessary for curvature in the response curves to be visible. Implementation of these two suggestions would allow for subsequent experimental setups having more success and allowing for more flexibility.

ii. Future Work

Should a successful experimental setup for equilibrium studies be built within *Imperial College London*, future work can commence by exploiting the flexibility of the ZLC technique. For equilibrium studies, experiments at different temperatures would allow for determination of the temperature dependence of the Henry constant; whilst experiments with different adsorbents would allow for comparative studies in relation to amount adsorbed. Experiments with the sorbate gas being a gas mixture can also be carried out - particularly important for carbon capture due to the nature of flue gas streams. Lastly, by employing the same experimental setup whilst altering the experimental conditions, such as experiments at higher purge flowrates, kinetic adsorption parameters can be easily extracted.

Acknowledgement

We would like to express our gratitude to Humera Ansari for her invaluable support in this project. Her help in building the experimental setup and provision of volumetric isotherm data was pivotal in the success of this research.

Nomenclature

β_n	Dimensionless parameter	
b	Langmuir 'b' constant	$\mathrm{bar}^{-1}/\mathrm{mL\,mol}^{-1}$
c	Sorbate gas concentration	$\mathrm{mol\,mL}^{-1}$
c_0	Initial sorbate gas concentration	$\mathrm{mol\,mL}^{-1}$
D	Diffusivity	$\mathrm{m}^2\,\mathrm{s}^{-1}$
F	Purge gas flowrate	$\mathrm{mL\,s}^{-1}$
K	Henry constant	$\mathrm{mmol\,g}^{-1}\,\mathrm{bar}^{-1}/\mathrm{mL}^{-1}$
L	Dimensionless parameter	
q	Adsorbed phase concentration	$\mathrm{mol\,mL}^{-1}$
t	Time	s
V_g	Dead volume of column	mL
V_s	Volume of adsorbent sample	mL
y_0	Initial sorbate gas mole fraction	

References

Brandani, F. et al. (2003). "Measurement of Adsorption Equilibrium by the Zero Length Column (ZLC) Technique Part 1: Single-Component Systems". In: pp. 1451–1461. DOI: 10.1021/ie020572n.

Brandani, S. (1996). "Analytical Solution for ZLC Desorption Curves with Bi-Porous Adsorbent Particles". In: 51.12.

– (1998a). "Effects of nonlinear equilibrium on zero length column experiments". In: 53.15, pp. 2791–2798.

– (1998b). "Heat Effects in ZLC Experiments". In: 285, pp. 275–285.

Brandani, S. et al. (1995). "Analysis of Desorption Curves for Liquid Systems". In: *Chemical Engineering Science* 50.13, pp. 2055–2059.

Brandani, S. et al. (1996). "Moments Analysis of the Zero Length Column Method". In: 7, pp. 315–319. DOI: 10.1021/ie950287m.

Brandani, S. et al. (2000). "Diffusion, self-diffusion and counter-diffusion of benzene and and p-xylene in silicalite". In: 36, pp. 283–300.

Clarivate Analytics (2019). *Web of science*. URL: https://apps.webofknowledge.com/.

Duncan, W. L. et al. (2002). "The effect of a crystal size distribution on ZLC experiments". In: 57, pp. 2641–2652.

Eberhart, C et al. (2000). "Comparing Inertia Weights and Constriction Factors in Particle Swarm Optimization". In: 7, pp. 84–88.

Friedrich, D. et al. (2015). "Automatic estimation of kinetic and isotherm parameters from ZLC experiments". In: 126, pp. 616–624. DOI: 10.1016/j.ces.2014.12.062.

Gelles, T. et al. (2020). "Separation and Purification Technology Diffusion kinetics of ethane, ethylene, and their binary mixtures in ethane-selective adsorbents". In: *Separation and Purification Technology* 230.May 2019, p. 115872. ISSN: 1383-5866. DOI: 10.1016/j.seppur.2019.115872.

Grenier, P. et al. (1994). "Diffusion of methanol in NaX crystals : Comparison of i.r., ZLC, and PFG-n.m.r. measurements". In: 14.

Gunadi, A. et al. (2006). "Diffusion of linear paraffins in NaCaA studied by the ZLC method". In: 90, pp. 278–283. DOI: 10.1016/j.micromeso.2005.11.027.

Haupt, R. L. et al. (2003). *Practical Genetic Algorithms*. 2nd ed. John Wiley & Sons, p. 261. ISBN: 9780471455653.

Hu, X. et al. (2014). "Diffusion mechanism of CO_2 in 13X zeolite beads". In: pp. 121–135. DOI: 10.1007/s10450-013-9554-z.

Hu, X. et al. (2015a). "Development of a Semiautomated Zero Length Column Technique for Carbon Capture Applications: Rapid Capacity Ranking of Novel Adsorbents". In: DOI: 10.1021/acs.iecr.5b00513.

– (2015b). "Development of a Semiautomated Zero Length Column Technique for Carbon Capture Applications: Study of Diffusion Behavior of CO_2 in MOFs". In: DOI: 10.1021/acs.iecr.5b00515.

Keller, J. et al. (2005). *Gas Adsorption Equilibria: Experimental Methods and Adsorption Isotherms*. Springer. ISBN: 0-387-23598-1.

Krishnamurthy, S et al. (2019). "Adsorption and diffusion of CO_2 in CPO-27–Ni beads". In: *Adsorption* August. ISSN: 1572-8757. DOI: 10.1007/s10450-019-00162-x.

Kumar, V. et al. (2018). "Comparative Studies on CO_2 Adsorption Isotherms by Solid Adsorbents". In: *Materials Today: Proceedings* 5.11, pp. 23033–23042. ISSN: 2214-7853. DOI: 10.1016/j.matpr.2018.11.032.

Mangano, E et al. (2013a). "Efficient and Rapid Screening of Novel Adsorbents for Carbon Capture

in the UK IGSCC Project". In: *Energy Procedia* 37.0, pp. 40–47. ISSN: 1876-6102. DOI: 10.1016/j.egypro.2013.05.083.

Mangano, E. et al. (2013b). "Analysis and Interpretation of Zero Length Column Response Curves". In: 11, pp. 1714–1718. DOI: 10.1002/cite.201300083.

Mladen, E. et al. (1988). "A new experimental technique for measurement of intracrystalline diffusivity". In: 8, pp. 40–45.

Ruthven, D. M. (1984). *Principles of Adsorption and Adsorption Processes.*

Vidoni, A. et al. (2012). "Diffusion of C_2H_6 and C_2H_4 in DDR Zeolite". In: pp. 1383–1390. DOI: 10.1021/ie202449q.

Ceramic Hollow Fibre Supported ZIF-7 Membrane for Hydrogen Separation

Weixuan Kong and Zhaomin Yu
Department of Chemical Engineering, Imperial College London, U.K.

Abstract H_2 selective membrane is a promising substitute for the conventional hydrogen separation techniques. Among various types of the membrane being studied, zeolitic-imidazolate frameworks, in particular, ZIF-7 membrane emerged as a novel candidate for hydrogen separation. In this study, ZIF-7 membrane was fabricated at the outer surface of ceramic hollow fibre by *in-situ* crystallisation at 105°C with various synthesis time 1h, 2h and 3h. The 3-hour synthesised membrane was well-intergrown and defect-free, with a thickness of around 2μm at the outer surface of the hollow fibre. The gas permeation tests showed that an H_2/N_2 separation factor of 7.41 and an H_2/CO_2 separation factor of 14.4 were obtained, exceeding the H_2/N_2 (3.7) and H_2/CO_2 (4.7) Knudsen separation factors respectively. It exhibits better separation results than some of the ZIF-7 membranes with alumina disk-shaped support and alumina hollow fibre support synthesised from a different method.

Introduction

Hydrogen element exists in greatest amount in the universe. It is unscented and nontoxic. It possesses the highest energy content by weight of the ordinary fuels - almost three times that of gasoline (Valladares, 2017). It is thus proposed as a promising renewable energy resource, and it offers reliable, sustainable, environmental and climate-friendly power with a high-energy density that meets global energy requirements (Chang *et al.*, 2018a). Hydrogen is also an "energy carrier" like electricity. It is commonly applied and used in all sectors of the economy: power, transportation, industry and buildings. However, it is typically produced together with other gases (CO_2, N_2, CH_4, etc.) during industrial processes such as steam reforming of hydrocarbons with water-gas shift (WGS) (Nikolaidis & Poullikkas, 2017; Qiu *et al.*, 2014). Hydrogen separation thus becomes an important industrial process to obtain the high-purity hydrogen stream as the main product. In this project, it is aimed to explore a simple and efficient method for hydrogen separation.

Background

The conventional hydrogen separation methods used commercially include pressure swing adsorption (PSA) and cryogenic distillation. These two techniques are energy-intensive and produce relatively low-purity hydrogen. H_2 selective membrane has been proposed as an substitute for high-purity hydrogen separation due to its ease of operation, low energy requirement and investment cost (Adhikari & Fernando, 2006; Li, *et al.*, 2010a). In order to be an effective substitute, it was suggested that the minimum H_2/CO_2 separation factor needs to exceed the Knudsen separation factor of 4.7 (Go *et al.,* 2016). Several studies have been exploring the use of inorganic membranes in the separation. Among those, palladium (Pd)-based membrane was the one being heavily studied (Adhikari & Fernando, 2006). Theoretically, Pd-based membrane possesses extremely high H_2/CO_2 separation factor and satisfactory performance at a temperature higher than 500 °C (Li *et al.,* 2010a). However, the performance of the membrane would be drastically reduced due to the surface contamination while contacting with various impurities such as CO and H_2S (Adhikari & Fernando, 2006). Another barrier to the large-scale use of Pd-based membrane is attributed to the expensive Pd cost (Li *et al*., 2010a).

Another type of membrane being studied is the carbon molecular sieve (CMS). It shows promising performance with high-purity feed (Li *et al*., 2010a). However, the brittleness of the material hinders the commercialization in large-scale separation process (Koros & Mahajan, 2000). The membrane also needs to be supported to be mechanically stable in industrial applications (Briceño *et al*., 2012). A few studies on fabrication of supported membrane in the membrane reactor have been conducted successfully. Nevertheless, the problem of obtaining a defect-free membrane while coating the inorganic layers on the porous surfaces has been encountered in those studies (Briceño *et al*., 2012). Although several strategies have been proposed to resolve the problem, those techniques increase the cost and the synthesis time (Briceño *et al.*, 2012).

Microporous amorphous silica membrane is also being explored as a possible separation candidate. Although it is able to exhibit outstanding H_2 separation performance over other smaller-sized molecular gases, the membrane suffers from material instability in the humid environment (Lin *et al.,* 2002). Moreover, the relatively larger zeolitic pore size of the highly siliceous microporous zeolite membrane compared to the kinetic diameters of both CO_2 and H_2, together with the inter-crystalline micro defects, contribute to the excellent hydrothermal stability but low H_2/CO_2 separation factor (Gu, Tang & Dong, 2008; Hong, Falconer & Noble, 2005; Kanezashi *et al.,* 2008). Therefore, it is necessary to introduce post-treatment improvement as well as pore size engineering to address these issues (Gu, Tang & Dong, 2008).

Metal-organic frameworks (MOFs) have emerged as the latest category of nanoporous crys-

talline materials. They have a high surface area, regular and highly tunable pore size, and rich physiochemical properties that enable them to be the promising candidates for storage and separation of gases (Liu *et al.,* 2017). Zeolitic-imidazolate frameworks (ZIFs) are a sub-class of MOFs. They have become new candidates for the use in gas and vapour separations. They are composed of metals with tetrahe-

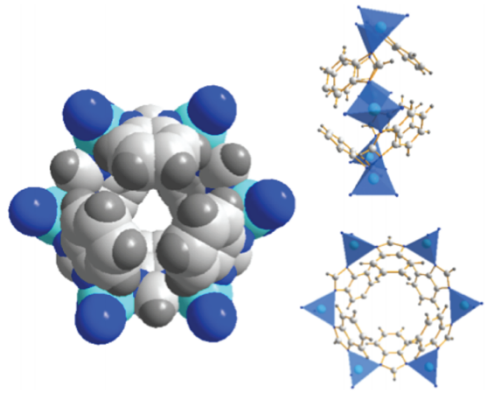


Figure 1. Left part shows the main cavity of ZIF-7. Top-right and bottom show lateral and front views of the six-membered-ring (6MR) pore opening respectively. Zinc clusters are represented as polyhedral (McCarthy *et al*., 2010).

dral coordination geometry and imidazole ligands, forming a zeolite-like structure with permanent porosity, uniform pore size, and exceeding thermal and chemical stability (Park *et al*., 2006). Among the various ZIFs, ZIF-7 membranes are found achieving a high separation factor of H_2 over CO_2 according to Li *et al* (2010a), through a molecular sieving effect (Li *et al.,* 2010a). ZIF-7 has an open-framework structure with sodalite (SOD) topology in hexagonal symmetry as shown in Figure 1, formed by connecting benzimidazolate (BIM) anions and zinc cations (Huang, Zhang & Chen, 2003). The size of H_2 is 0.29 nm and that of CO_2 is 0.33 nm while ZIF-7 has a pore size just in between them (~0.3 nm), enabling membrane molecular sieving separation for H_2/CO_2 mixture (Li, *et al.*, 2010a).

There have been several investigations regarding the different fabrication methods of ZIF-7 membranes. *In-situ* crystallization is regarded as the most common technique used for zeolite membrane synthesis. In this method, the support is placed and contacted the precursor solutions directly, enabling the heterogeneous nucleation and production of well-intergrown membrane layer on the support surface (Ma & Liu, 2019). McCarthy *et al.* (2010) first synthesised continuous crack-free ZIF-7 membranes on alumina support with 1 N, N dimethylformamide (DMF) as solvent by *in-situ* technique. This method is simple and does not require seeding or surface modification. The existence of sodium formate was found to be essential for one-step *in-situ* fabrication of continuous ZIF-7 membranes layer. Using the same synthesis conditions without the sodium formate in the growth solution, it was observed that considerably less ZIF-7 nucleated on the support (McCarthy *et al*., 2010). Some other reports also explored the synthesis of ZIF-7 in other solvents such as water-ethanol mixtures (He *et al*., 2013) and N, N-diethylformamide (Seoane *et al*., 2012).

Another technique for ZIF-7 membrane preparation is the secondary seeded growth method. ZIF-7 membranes were firstly synthesised on alumina supports by Li *et al.* (2010a) using a microwave-assisted solvothermal secondary seeded growth method. By separating the nucleation and crystal growth steps, a better command of membrane formation and quality can be obtained. In this technique, ZIF-7 crystals are firstly synthesised and coated on the support surface at room temperature, hydrothermal or solvothermal synthesis is then followed to grow a continuous thin film (Ma & Liu, 2019). Li *et al.* (2010a) used a dip-coating technique for the surface seeding of the alumina support. Although this method is reported to produce a very thin ZIF-7 top layer (Li *et al.,* 2010a), the synthesised ZIF-7 membranes usually contain guest molecules in their cavities that need to be removed with methanol or treated at a high temperature (Gücüyener *et al*., 2010a; He *et al*., 2013).

One distinct property about ZIFs synthesis to be noted is the rapid reaction within the metal and ligand precursors in the absence of external structure-

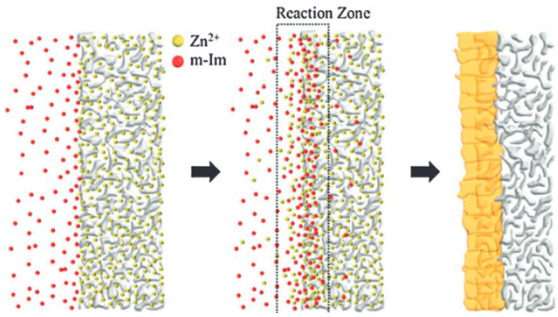


Figure 2. Counter-diffusion technique illustration (Kwon & Jeong, 2013)

directing agents (SDA). Taking this advantage, innovative ideas to synthesise membranes can therefore be implemented to boost the separation performance (Ma & Liu, 2019). One noteworthy approach is based on the theory of counter-diffusion firstly developed by Yao *et al.* (2011). A "reaction zone" is formed at the interface by the diffusion of metal ions and ligand molecules as shown in Figure 2, rapid heterogeneous nucleation then takes place at this location which leads to the continuous crack-free ZIF membranes (Kwon & Jeong, 2013). This approach has been further improved to an *in-situ* fabrication method based on the counter-diffusion concept by Jeong *et al.* (2013). In addition, although Jeong *et al.* only synthesised ZIF-8 from this method, it is reported that this method is able to heal the defects in the membrane and it is capable to synthesise ZIF-7 membranes as well (Ma & Liu, 2019).

Recently, Kim *et al.* also proposed an innovative

direct spraying technique for continuous ZIF membranes fabrication that is simple and time-efficient and operates at the ambient pressure (Melgar, Kwon & Kim, 2014). By spraying with the aid of electrostatic force, fine liquid droplets of the precursors are deposited on the heated alumina supports surface and the membrane layer is formed directly as shown in Figure 3. The membrane thickness can be varied between 2 µm and 22 µm by modifying the sprayed volume of the ZIFs precursors and the deposition time. Using this approach, the H_2/CO_2 separation factor and H_2 permeance were reported to be 9.6 and

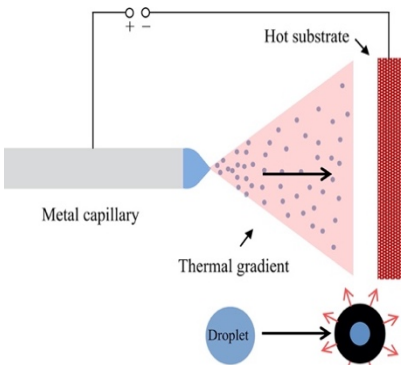


Figure 3. Direct spraying technique illustration (Melgar *et al.* 2014)

4.61×10^{-7} mol·Pa^{-1} ·m^{-2} ·s^{-1} respectively for the ZIF-7 membrane at 25 °C.

The impact on separation performance with various membrane configurations has also been investigated in a range of studies. The majority of the investigation is on the α-Al_2O_3 disk-shaped support with different synthesis conditions (Melgar, Kim & Othman, 2015). Among those, the highest H_2/CO_2 separation factor was obtained to be 15.98 using the *in-situ* growth method reported by Kim *et al* (2016). However, this relatively high separation factor came at the expense of the H_2 permeability (Go *et al*., 2016). Moreover, compared with ZIF-7 membrane performance using the Monte Carlo molecular simulation (Thornton *et al.*, 2012), this method failed to achieve the industrial requirement in terms of selectivity (Go *et al.,* 2016). Further improvement is required to operate commercially. Alternatively, the study on ZIF-7 membrane with α-Al_2O_3 hollow fibre support under solvothermal fabrication was suggested by Yao *et al* (2013). It was reported that the initial H_2/N_2 selectivity (2.7) failed to exceed that of Knudsen diffusion (~3.7). A post-modification was required to tune the H_2/N_2 selectivity into 3.9, right above the Knudsen diffusion selectivity. Additionally, ceramic is an unparalleled material because of its high chemical, thermal and mechanical resistance, allowing operation under a harsh condition such as the use of an organic solvent (Jiang *et al*., 2013). Therefore, it shows a large room of improvement for using the ceramic hollow fibre configuration.

It is proposed that membranes produced from *in-situ* fabrication typically exhibit larger thickness and this technique is commonly used for hydrogen separations (Ma & Liu, 2019). Compared with other crystallisation methods, *in-situ* fabrication is also the simplest and fastest. Therefore, in this study, it is decided that ZIF-7 membrane fabrication will be conducted using the *in-situ* method, and ZIF-7 membranes will be produced with minimal defects with ceramic hollow fibre support under various synthesis conditions. Different characterisations for the membrane will be investigated and the effective separation performance for H_2/CO_2 and H_2/N_2 mixture gases will also be examined.

Methods

ZIF-7 Membrane Synthesis Method

ZIF-7 membrane was prepared by immersing the clean ceramic hollow fibre into the precursor solution. The crystals formed on the outer surface of the hollow fibre under the temperature of 105°C with different synthesis time (1h, 2h and 3h). The ceramic hollow fibre (~3.5cm in length) was firstly sonicated with DI water and acetone sequentially. 7.65 g zinc nitrate hexahydrate (99%, Sigma-Aldrich) and 4.05 g benzimidazole (98%, Aldrich) were dissolved in 100 mL dimethylformamide (DMF) (VWR Chemicals) and stirred for 30 min. Another separate solution was prepared by dissolution of 1.40 g sodium formate (99%, Sigma) into 100 mL DMF with stirring for 30 min as well. These two solutions were then mixed and stirred for another 10 min. Ceramic hollow fibre was wrapped with Teflon tape on both ends, plugged into a Teflon-lined stand and placed vertically into a glass vial. The mixture solution was transferred into the vial with hollow fibre inside. After crystallisation, the hollow fibre was cooled to room temperature and washed with methanol (VWR Chemicals) twice with the first wash left overnight and second wash for 5 min. The hollow fibre was then left to dry under the saturated condition at room temperature for 2 days.

Characterisation Methods

The dried ZIF-7 powder was characterised with X-Ray Diffraction (XRD) and scanning electron microscope (SEM). XRD patterns were measured for the ZIF-7 particles on a PANalytical X'Pert Pro in the range of 5-45° with a step size of 0.02°. Additionally, the SEM images were examined on a ZEISS electron microscope. The accelerating voltage and working distance were set 5 kV and 4-6 mm. Before using the SEM, the membrane was first fractured, cleaned with the airflow to remove dirt and coated with a nanometer-scale gold layer which helps to reduce the charging effects (Li *et al*., 2010a).

Gas Permeation Test

All gas permeation tests were conducted using the Wicker-Kallenbach technique through an online gas chromatography machine as shown in Figure 4 for single gas separation and Figure 5 for binary gas

separation. The single gases (H_2, N_2, CO_2 and CH_4) or a mixture of gases (H_2/CO_2, H_2/N_2) were fed into the shell side while the sweep gas flows through the lumen side.

For the single gas permeation, the feed flow rate was kept as 30 mlmin^{-1}. The pressures on both feed and permeate sides were set at 1 bar. The single gas permeation was conducted for comparison purpose.

In the mixed gas permeation tests, the total feed flow rate was set at 60 mlmin^{-1} where each component was 30 mlmin^{-1}. The pressures on both sides were also kept stable at 1 bar.

For all measurements, argon was used as the sweep gas and the flow rate was set at 15 mlmin^{-1}.

To ensure the reliability of gas tests, three membrane samples synthesised under the same condition were used for repeated measurements.

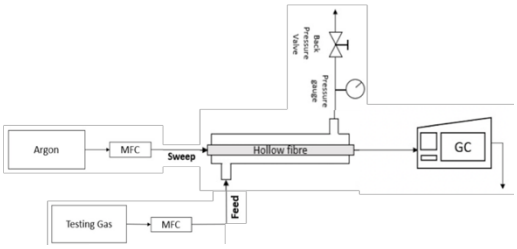


Figure 4. Single gas permeation illustration

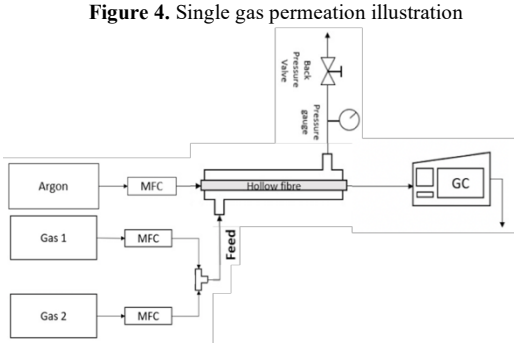


Figure 5. Binary gas permeation illustration

The membrane permeance, F_i, is calculated using:

$$F_i = \frac{N_i}{\Delta P_i A} \quad (1)$$

where N_i is the permeate molar rate of component i (mols^{-1}), ΔP_i is the pressure drop of i across the membrane (Pa), and A is the surface area of membrane (m^2).

For the separation factor, $\alpha_{H_2,j}$, it is defined as:

$$\alpha_{H_2,j} = \frac{y_{H_2}/y_j}{x_{H_2}/x_j} \quad (2)$$

where x and y are the mole fractions in the retentate and permeate respectively, j represents CO_2 or N_2.

The permeate flow rate across the membrane can be assumed to be negligible compared with that of the feed and the sweep gas. Thus, the component concentrations in the retentate and the feed are almost identical. Since 1:1 mixture was decided for feed flow rate, the equation can be simplified to:

$$\alpha_{H_2,j} = \frac{y_{H_2}}{y_j} \quad (3)$$

According to the definition for mole fraction:

$$y_{H_2} = \frac{permeate\ flow\ rate\ of\ H_2}{Overall\ permeate\ flow\ rate} \quad (4)$$

and $y_j = \frac{permeate\ flow\ rate\ of\ j}{Overall\ permeate\ flow\ rate}$, (5)

the separation factor can be further simplified to the ratio of the permeate flow rate of H_2 over that of j, which is equivalent to the equation of selectivity.

Results and Discussions

Characterisation of ZIF-7 Membrane

The as-synthesised ZIF-7 particles were characterised by XRD. Figure 6a indicates that there are consistent XRD patterns for both the particles synthesised at 1h, 2h and 3h. Compared with the location of the main peaks in both simulated and experimental literature value in Figure 6b, it can be confirmed that the correct pure ZIF-7 particles were synthesised without ZnO impurities. It is noticeable that the XRD pattern for 2-hour synthesised particle exhibits lower intensity at all peaks. This may be attributed to the preferential orientation of the particle structure. The drying method and conditions to form

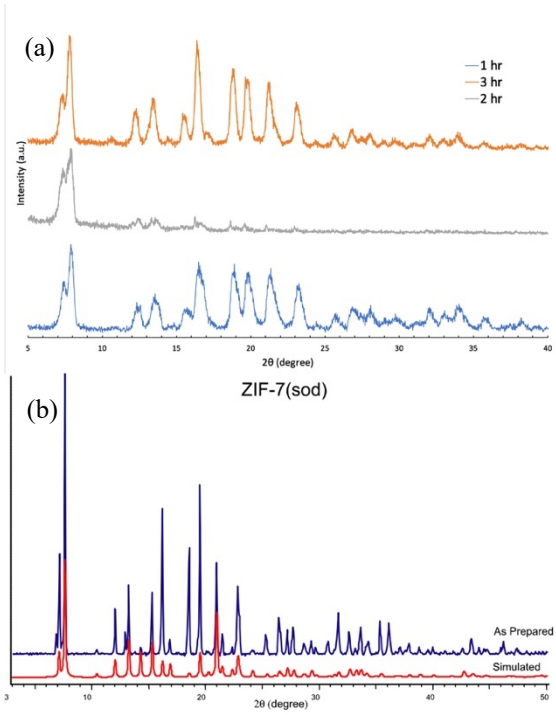


Figure 6. (a) experimental XRD pattern for ZIF-7 particle, and (b) Reference XRD pattern for ZIF-7 particle (Park *et al*., 2006b)

powder may affect the preferential orientation. More importantly, upon drying, some particles may orient in specific directions, resulting in non-spherical particle shapes. The preferential orientation would lead to a systematic error in the diffraction peak intensities (Holder & Schaak, 2019). Repetition of this measurement may be helpful to confirm the XRD result for 2-hour synthesised ZIF-7 particle. Moreover, it can be observed that the peaks for 3-hour synthesised particles are sharper with higher intensities than that for 1-hour particles. The 1-hour synthesised ZIF-7 particles may contain some amorphous

components which lack long-range order, compared with the 3-hour synthesised particles. The X-ray, therefore, would be scattered in a range of direction, producing the broad scattering peaks (Holder & Schaak, 2019). It is presumed that the ZIF-7 particles have better-grown crystalline structure with prolonged synthesis time.

The ZIF-7 membranes morphologies were investigated using SEM. The SEM image illustrates that the ZIF-7 particles have a cubic or rectangular shape, agreed with the ZIF-7 particle morphology in the literature (Gücüyener *et al.* 2010a). A thin membrane layer of ZIF-7 membrane synthesised for 1 hour at the outer surface of the hollow fibre is observed with the thickness about 1 μm, as shown in Figure 7a. The inner surface view of the ceramic hollow fibre in Figure 7b indicates no penetration of the ZIF-7 particles through the hollow fibre pore into the inner surface since the pore size (~200nm) is much smaller than the size of the particle. From the zoomed-in top view of ZIF-7 membrane in Figure 7c, multiple inter-crystalline defects are observed. Therefore, it can be deduced that 1 hour is insufficient for the growth of continuous crack-free ZIF-7 membrane.

The 2-hour synthesised membrane gave a similar

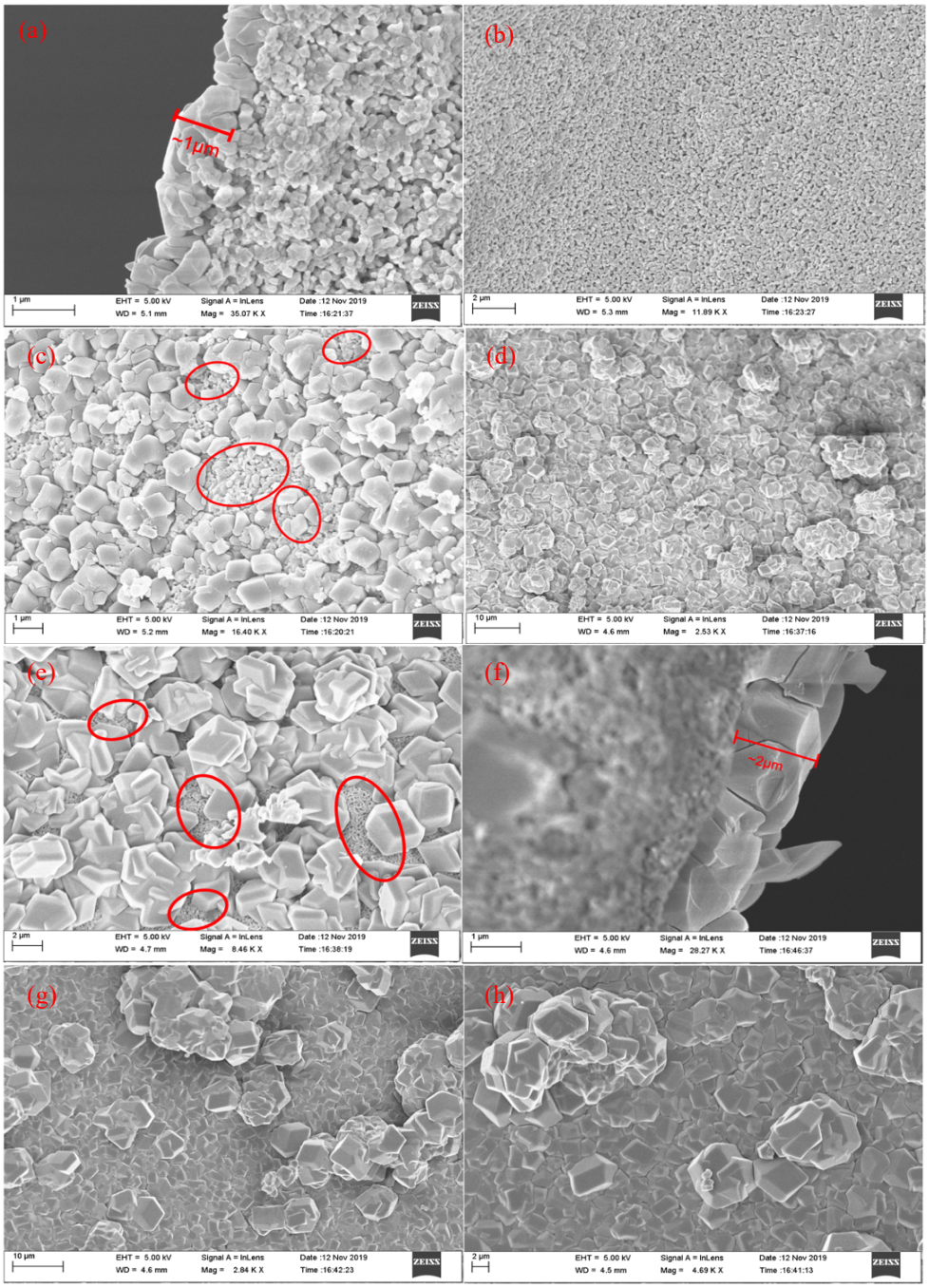


Figure 7. (a) 1h membrane cross-sectional view, (b) inner view of hollow fibre for 1h membrane, (c) 1h membrane top view, (d) 2h membrane top view at 10μm scale, (e) 2h membrane top view at 2 μm scale, (f) 3h membrane cross-sectional view, (g) 3h membrane top view at 10 μm scale and (h) 3h membrane top view at 2 μm scale

result to the 1-hour one. Figure 7d and 7f show the top view of the ZIF-7 membrane synthesised in 2 hours. Under a 10 μm scale, a continuous layer is seemed to form. While zooming into a 2 μm scale, there are still several defects present between crystals, although the defects are fewer and smaller than those observed in the 1-hour membrane. Both membranes synthesised under the temperature of 105°C for 1 hour and 2 hours appeared to be defective, it is suggested that H_2 gas molecules tend to permeate through the ZIF-7 membrane via simple diffusion rather than membrane separation (molecular sieving).

The cross-sectional view in Figure 7f exhibits that ZIF-7 membrane synthesised for 3 hours has a thickness of around 2 μm, which is still quite thin compared to MOF membranes synthesised from other literatures (Arnold *et al.*, 2007; Gascon, Aguado & Kapteijn, 2008; Liu, Y. *et al.*, 2009; Yoo, Lai & Jeong, 2009; Yao *et al.* 2013). Figure 7g and 7h illustrate the top view of membranes under 10 μm scale and 2 μm scale respectively. The two SEM images both show a continuous crack-free membrane layer. Generally, the ZIF-7 membrane grows thicker with prolonged crystallisation time, and at least 3 hours are required for a well-intergrown continuous ZIF-7 membrane.

A thin membrane is critical to achieve high permeances (Li *et al.*, 2010a). It was reported that a thicker membrane results in lower H_2 permeance in the literature (Brunetti *et al.*, 2017). This may be explained by Darcy's Law:

$$Q = -\frac{kA(p_b - p_a)}{\mu L} \quad (6)$$

where Q is the discharge flow rate through the membrane, k represents the intrinsic permeability of the membrane, A as the cross-sectional area of the membrane, (p_b-p_a) is the pressure drop across the membrane, μ as the dynamic viscosity and L represents the thickness of the membrane. With a negligible change in pressure drop and cross-sectional area after membrane formation, an increase in thickness leads to a reduction in permeate flowrate, lowering the permeance according to Equation 1. Therefore, it is worthwhile to optimise the thickness with desirable permeance in the future study.

Gas Permeation Performance of ZIF-7 Membrane

The gas permeation test performance of equimolar H_2/N_2 mixtures and H_2/CO_2 mixtures using ZIF-7 membranes formed on ceramic hollow fibre outer surface were tested at room temperature. Table 1 gives the permeance and separation factor of membranes synthesised with different conditions.

It is clear that there is no H_2/N_2 separation performance obtained for ZIF-7 membranes synthesised under 105°C for 1 hour and 2 hours through molecular sieving, and this coincides with the SEM results illustrated before which shows a defective membrane. However, the ZIF-7 membrane synthesised under the same temperature in 3 hours exhibits performance in both H_2/N_2 and H_2/CO_2 mixtures. The separation factors for the equimolar H_2/N_2 and H_2/CO_2 mixtures are 7.41 and 14.4 respectively, both are significantly greater than the matching Knudsen separation factors which are 3.7 and 4.7. This confirms that the synthesised ZIF-7 membrane in this project contributes effectively to the separation of gases (Farjoo *et al.*, 2017).

ZIF-7 membranes synthesised on ceramic hollow fibre in this study exhibit a much higher separation factor (7.41) for H_2/N_2 mixtures than that of the α-Al_2O_3 hollow fibre (3.9) reported by Yao *et al.* (2013). The synthesis time used in this study (3h) is also considerably shorter than that used by Yao *et al.* (8h). This enhancement in synthesis procedure and separation performance is attributed to the improved structure of the ceramic hollow fibre, possibly due to the elimination of the redundant dense surface and the sponge-like layer which creates additional mass transfer resistance (Lee, Wang & Li, 2016). A higher separation factor is also obtained for H_2/CO_2 mixtures with ZIF-7 membranes on ceramic hollow fibre than that of alumina disk used by Li et al. (Li *et al.*, 2010a). This improved separation factor is probably because of the high surface area of hollow fibre supports and the complete elimination of guest molecules inside the ZIF-7 framework.

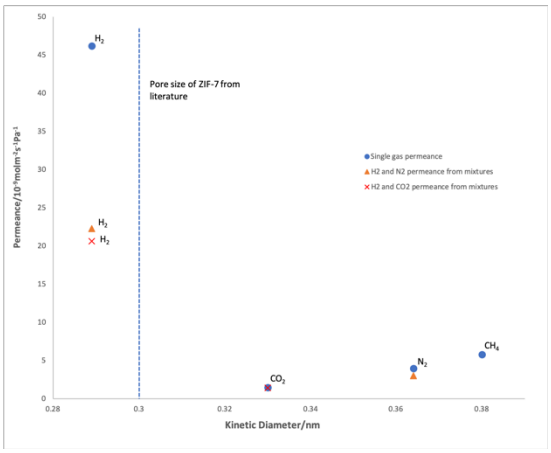


Figure 8. Single and mixed (1:1 mixtures) gas permeances measured from ZIF-7 membranes on ceramic hollow fibre synthesised under 105°C 3 hours as a function of molecular kinetic diameters

Single gas permeation performances of H_2, CO_2, N_2 and CH_4 were obtained using ZIF-7 membranes synthesised from 105°C and 3 hours for comparison purpose. Figure 8 shows the single and mixed gas permeances against the kinetic diameter of the gas molecules. The permeance of single gases generally follows the order of their molecular sizes. An apparent cut-off between H_2 and other gas molecules that were greater than the ZIF-7 pore size (0.3 nm) can be observed (Gücüyener *et al.*, 2010a). Also, it is notable that the permeances are almost the same for CO_2 and N_2 both as single gases and from the equimolar binary mixtures. However, the permeance

for H_2 as single gas ($46.2\times10^{-9}molm^{-2}s^{-1}Pa^{-1}$) is much greater than that of H_2 from the equimolar mixtures ($22.3\times10^{-9}molm^{-2}s^{-1}Pa^{-1}$ for H_2/N_2 and $20.6\times10^{-9}molm^{-2}s^{-1}Pa^{-1}$ for H_2/CO_2, respectively). This might result from the competition between the adsorption of CO_2 and N_2 gas molecules onto the membrane pore surface and the molecular sieving diffusion of H_2 gas molecules through the membrane in the mixtures (Farjoo *et al*., 2017). Thus, diffusion of H_2 gas molecules into pores of ZIF-7 membrane is hindered.

The ideal separation factor of H_2 over CO_2, N_2 and CH_4 are 11.7, 31.9 and 7.95, respectively as shown in Table 2. These values are all greater than the corresponding Knudsen separation factors (3.7, 4.7 and 2.8, separately), demonstrating the high compactness of the synthesised membrane (Chang *et al*., 2018). It can be observed that the actual separation factors of equimolar H_2/N_2 and H_2/CO_2 binary mixtures (7.41 and 14.4, respectively) deviate from the ideal separation factors (11.7 and 31.9, respectively). This is attributed to the strong interaction within different molecules and with the porous system of the membrane (He, Z. & Wang, 2018). In the case of zeolite membranes, this deviation is often observed in many other literatures as well (Gu, Tang & Dong, 2008; Guan *et al*., 2003; Li *et al*., 2010a).

Table 1 Gas permeation results under different synthesis conditions

ZIF-7 Ceramic Hollow Fibre Membranes	Permeance ($\times10^{-9}molm^{-2}s^{-1}Pa^{-1}$)				Separation Factor	
Synthesis Conditions	H_2	N_2	H_2	CO_2	H_2/N_2	H_2/CO_2
105°C 1 hour	7.67	N/A	-	-	N/A	-
105°C 2 hours	1.11	N/A	-	-	N/A	-
105°C 3 hours	22.3	3.01	20.6	1.43	7.41	14.4

Table 2 Single gas permeation results under 105°C 3 hours

Single gas permeation using membranes synthesised from 105°C 3 hours	Permeance ($\times10^{-9}molm^{-2}s^{-1}Pa^{-1}$)	Ideal Separation Factor	Knudsen Separation Factor
H_2	46.2	-	-
N_2	3.95	11.7	3.7
CO_2	1.45	31.9	4.7
CH_4	5.81	7.95	2.8

Conclusions

The continuous, defect-free ZIF-7 membrane with a thickness of 2 μm was successfully synthesised by one-step *in-situ* crystallisation with ceramic hollow fibre at 105°C in 3 hours, characterised with XRD and SEM, and evaluated using single and binary gas permeation tests. The H_2/N_2 separation factor of 7.41 and H_2/CO_2 separation factor of 14.4 were obtained. In order to achieve better separation performance, post-modifications such as using β-cyclodextrin solution is a possible approach to reduce the inter-crystalline defects and enhance the H_2/N_2 separation factor. Besides, the further examination on how temperature would affect the separation factors should also be conducted, optimising the combination of temperature and synthesis time. The hydrothermal stability of the synthesised ZIF-7 membrane can also be tested to show the effectiveness of this separation method as a promising candidate for hydrogen separation for industry uses. To improve H_2 permeance, the thickness of the membrane should be taken into consideration. A well-intergrown thin membrane should be used while it is also worth noting the trade-off between selectivity and permeability with changing membrane thickness.

Acknowledgements

Weixuan Kong and Zhaomin Yu would like to thank Tongrong Wu and Dr. Nicholaus Prasetya for the technical support during the research project.

References

Adhikari, S. & Fernando, S. (2006) Hydrogen Membrane Separation Techniques. *Industrial & Engineering Chemistry Research.* 45 (3), 875-881. Available from: http://dx.doi.org/10.1021/ie050644l. Available from: doi: 10.1021/ie050644l.

Arnold, M., Kortunov, P., Jones, D. J., Nedellec, Y., Kärger, J. & Caro, J. (2007) Oriented Crystallisation on Supports and Anisotropic Mass Transport of the Metal-Organic Framework Manganese Formate. *European Journal of Inorganic Chemistry.* 2007 (1), 60-64. Available from:

https://doi.org/10.1002/ejic.200600698. Available from: doi: 10.1002/ejic.200600698.
Briceño, K., Iulianelli, A., Montané, D., Garcia-Valls, R. & Basile, A. (2012) *Carbon molecular sieve membranes supported on non-modified ceramic tubes for hydrogen separation in membrane reactors.* Available from: http://www.sciencedirect.com/science/article/pii/S0360319912014528.
Chang, H., Wang, Y., Xiang, L., Liu, D., Wang, C. & Pan, Y. (2018a) *Improved H2/CO2 separation performance on mixed-linker ZIF-7 polycrystalline membranes.* Available from: http://www.sciencedirect.com/science/article/pii/S0009250918304962.
Chang, H., Wang, Y., Xiang, L., Liu, D., Wang, C. & Pan, Y. (2018b) *Improved H2/CO2 separation performance on mixed-linker ZIF-7 polycrystalline membranes.* Available from: http://www.sciencedirect.com/science/article/pii/S0009250918304962.
Gascon, J., Aguado, S. & Kapteijn, F. (2008) *Manufacture of dense coatings of Cu3(BTC)2 (HKUST-1) on α-alumina.* Available from: http://www.sciencedirect.com/science/article/pii/S138718110700666X.
Go, Y., Lee, J. H., Shamsudin, I. K., Kim, J. & Othman, M. R. (2016) *Microporous ZIF-7 membranes prepared by in-situ growth method for hydrogen separation.* Available from: http://www.sciencedirect.com/science/article/pii/S0360319915023502.
Gu, X., Tang, Z. & Dong, J. (2008) *On-stream modification of MFI zeolite membranes for enhancing hydrogen separation at high temperature.* Available from: http://www.sciencedirect.com/science/article/pii/S1387181107004891.
Guan, G., Tanaka, T., Kusakabe, K., Sotowa, K. & Morooka, S. (2003) *Characterization of AlPO4-type molecular sieving membranes formed on a porous α-alumina tube.* Available from: http://www.sciencedirect.com/science/article/pii/S0376738802005446.
Gücüyener, C., van, d. B., Gascon, J. & Kapteijn, F. (2010a) Ethane/Ethene Separation Turned on Its Head: Selective Ethane Adsorption on the Metal−Organic Framework ZIF-7 through a Gate-Opening Mechanism. *Journal of the American Chemical Society.* 132 (50), 17704-17706. Available from: https://doi.org/10.1021/ja1089765. Available from: doi: 10.1021/ja1089765.
Gücüyener, C., van, d. B., Gascon, J. & Kapteijn, F. (2010b) Ethane/Ethene Separation Turned on Its Head: Selective Ethane Adsorption on the Metal−Organic Framework ZIF-7 through a Gate-Opening Mechanism. *Journal of the American Chemical Society.* 132 (50), 17704-17706. Available from: https://doi.org/10.1021/ja1089765. Available from: doi: 10.1021/ja1089765.
He, M., Yao, J., Li, L., Wang, K., Chen, F. & Wang, H. (2013) Synthesis of Zeolitic Imidazolate Framework-7 in a Water/Ethanol Mixture and Its Ethanol-Induced Reversible Phase Transition. *ChemPlusChem.* 78 (10), 1222-1225. Available from: https://doi.org/10.1002/cplu.201300193. Available from: doi: 10.1002/cplu.201300193.
He, Z. & Wang, K. (2018) The 'ideal selectivity' vs 'true selectivity' for permeation of gas mixture in nanoporous membranes. *IOP Conference Series: Materials Science and Engineering.* 323 012002. Available from: doi: 10.1088/1757-899X/323/1/012002.
Holder, C. F. & Schaak, R. E. (2019) Tutorial on Powder X-ray Diffraction for Characterizing Nanoscale Materials. *ACS Nano.* 13 (7), 7359-7365. Available from: https://doi.org/10.1021/acsnano.9b05157. Available from: doi: 10.1021/acsnano.9b05157.
Hong, M., Falconer, J. L. & Noble, R. D. (2005) Modification of Zeolite Membranes for H2 Separation by Catalytic Cracking of Methyldiethoxysilane. *Industrial & Engineering Chemistry Research.* 44 (11), 4035-4041. Available from: https://doi.org/10.1021/ie048739v. Available from: doi: 10.1021/ie048739v.
Huang, X., Zhang, J. & Chen, X. (2003) Zn(bim)2] · (H2O)1.67: A metal-organic open-framework with sodalite topology. *Chinese Science Bulletin.* 48 (15), 1531-1534. Available from: https://doi.org/10.1007/BF03183954. Available from: doi: 10.1007/BF03183954.
Jiang, H., Meng, L., Chen, R., Jin, W., Xing, W. & Xu, N. (2013) *Progress on Porous Ceramic Membrane Reactors for Heterogeneous Catalysis over Ultrafine and Nano-sized Catalysts.* Available from: http://www.sciencedirect.com/science/article/pii/S1004954113604607.
Qiu, S., Xue, M. & Zhu, G. (2014) Metal-organic framework membranes: from synthesis to separation application. *Chemical Society Reviews.* 43(16), 6116-6140. Available from: doi : 10.1039/C4CS0059A.
Kanezashi, M., O'Brien-Abraham, J., Lin, Y. S. & Suzuki, K. (2008) Gas permeation through DDR-type zeolite membranes at high temperatures. *AIChE Journal.* 54 (6), 1478-1486. Available from: https://doi.org/10.1002/aic.11457. Available from: doi: 10.1002/aic.11457.
Koros, W. J. & Mahajan, R. (2000) Pushing the limits on possibilities for large scale gas separation: which strategies. *Journal of Membrane Science.* 175(2), 181-196. Available from: https://doi.org/10.1016/S0376-7388(00)00418-X
Kwon, H. T. & Jeong, H. (2013) In Situ Synthesis of Thin Zeolitic–Imidazolate Framework ZIF-8 Membranes Exhibiting Exceptionally High Propylene/Propane Separation. *Journal of the American Chemical Society.* 135 (29), 10763-10768. Available from: https://doi.org/10.1021/ja403849c. Available from: doi: 10.1021/ja403849c.
Li, Y., Liang, F., Bux, H., Feldhoff, A., Yang, W. & Caro, J. (2010) Molecular Sieve Membrane: Sup-

ported Metal–Organic Framework with High Hydrogen Selectivity. *Angewandte Chemie International Edition.* 49 (3), 548-551. Available from: https://doi.org/10.1002/anie.200905645. Available from: doi: 10.1002/anie.200905645.

Li, Y., Liang, F., Bux, H., Yang, W. & Caro, J. (2010a) *Zeolitic imidazolate framework ZIF-7 based molecular sieve membrane for hydrogen separation.* Available from: http://www.sciencedirect.com/science/article/pii/S0376738810001961.

Li, Y., Liang, F., Bux, H., Yang, W. & Caro, J. (2010b) *Zeolitic imidazolate framework ZIF-7 based molecular sieve membrane for hydrogen separation.* Available from: http://www.sciencedirect.com/science/article/pii/S0376738810001961.

Lin, Y. S., Kumakiri, I., Nair, B. N. & Alsyouri, H. (2002) MICROPOROUS INORGANIC MEMBRANES. *Separation and Purification Methods.* 31 (2), 229-379. Available from: https://doi.org/10.1081/SPM-120017009. Available from: doi: 10.1081/SPM-120017009.

Liu, M., Wang, L., Zheng, X. & Xie, Z. (2017) Zirconium-Based Nanoscale Metal–Organic Framework/Poly(ε-caprolactone) Mixed-Matrix Membranes as Effective Antimicrobials. *ACS Applied Materials & Interfaces.* 9 (47), 41512-41520. Available from: https://doi.org/10.1021/acsami.7b15826. Available from: doi: 10.1021/acsami.7b15826.

Liu, Y., Ng, Z., Khan, E. A., Jeong, H., Ching, C. & Lai, Z. (2009) *Synthesis of continuous MOF-5 membranes on porous α-alumina substrates.* Available from: http://www.sciencedirect.com/science/article/pii/S1387181108004599.

Ma, X. & Liu, D. (2019) Zeolitic Imidazolate Framework Membranes for Light Olefin/Paraffin Separation. *Crystals.* 9(1), 14. Available from: https://doi.org/10.3390/cryst9010014

McCarthy, M. C., Varela-Guerrero, V., Barnett, G. V. & Jeong, H. (2010) Synthesis of Zeolitic Imidazolate Framework Films and Membranes with Controlled Microstructures. *Langmuir.* 26 (18), 14636-14641. Available from: https://doi.org/10.1021/la102409e. Available from: doi: 10.1021/la102409e.

Melgar, V. M., Kim, J. & Othman, M. R. (2015) *Zeolitic imidazolate framework membranes for gas separation: A review of synthesis methods and gas separation performance.* Available from: http://www.sciencedirect.com/science/article/pii/S1226086X15000787.

Melgar, V. M., Kwon, H. T. & Kim, J. (2014) *Direct spraying approach for synthesis of ZIF-7 membranes by electrospray deposition.* Available from: http://www.sciencedirect.com/science/article/pii/S0376738814001343.

Nikolaidis, P. & Poullikkas, A. (2017) *A comparative overview of hydrogen production processes.* Available from: http://www.sciencedirect.com/science/article/pii/S1364032116305366.

Park, K. S., Ni, Z., Côté, A. P., Choi, J. Y., Huang, R., Uribe-Romo, F., Chae, H. K., O'Keeffe, M. & Yaghi, O. M. (2006) Exceptional chemical and thermal stability of zeolitic imidazolate frameworks. *Proc Natl Acad Sci USA.* 103 (27), 10186. Available from: http://www.pnas.org/content/103/27/10186.abstract. Available from: doi: 10.1073/pnas.0602439103.

Seoane, B., Zamaro, J. M., Tellez, C. & Coronas, J. (2012) Sonocrystallization of zeolitic imidazolate frameworks (ZIF-7, ZIF-8, ZIF-11 and ZIF-20). *CrystEngComm.* 14 (9), 3103-3107. Available from: http://dx.doi.org/10.1039/C2CE06382D. Available from: doi: 10.1039/C2CE06382D.

Thornton, A. W., Dubbeldam, D., Liu, M. S., Ladewig, B. P., Hill, A. J. & Hill, M. R. (2012) Feasibility of zeolitic imidazolate framework membranes for clean energy applications. *Energy & Environmental Science.* 5 (6), 7637-7646. Available from: http://dx.doi.org/10.1039/C2EE21743K. Available from: doi: 10.1039/C2EE21743K.

Valladares, M. (2017) Global Trends and Outlook for Hydrogen. *IEA Hydrogen Technology Collaboration Program (TCP).* Available from: https://ieahydrogen.org/pdfs/Global-Outlook-and-Trends-for-Hydrogen_Dec2017_WEB.aspx

Yao, J., Dong, D., Li, D., He, L., Xu, G. & Wang, H. (2011) Contra-diffusion synthesis of ZIF-8 films on a polymer substrate. *Chemical Communications.* 47 (9), 2559-2561. Available from: http://dx.doi.org/10.1039/C0CC04734A. Available from: doi: 10.1039/C0CC04734A.

Yao, J., Li, D., Wang, K., He, L., Xu, G. & Wang, H. (2013) Alumina Hollow Fiber Supported ZIF-7 Membranes: Synthesis and Characterization. *Journal of Nanoscience and Nanotechnology.* 13 (2), 1431-1434. Available from: https://www.ncbi.nlm.nih.gov/pubmed/23646654. Available from: doi: 10.1166/jnn.2013.5991.

Yoo, Y., Lai, Z. & Jeong, H. (2009) *Fabrication of MOF-5 membranes using microwave-induced rapid seeding and solvothermal secondary growth.* Available from: http://www.sciencedirect.com/science/article/pii/S1387181109001668.

Ultra-Thin Cyclodextrin Membranes with Aligned Intrinsic Pores for Both Polar and Non-Polar Organic Solvent Nanofiltration

Tomas Juergensen and Thomas Azimrayat Andrews
Department of Chemical Engineering, Imperial College London, U.K.

December 12th, 2019

Abstract

Amine-functionalised α-cyclodextrin (α-CD) monomers were reacted with terephtaloyl chloride (TPC) to form a selective thin film of intrinsic porosity. Macromolecules of intrinsic porosity have been implemented in the field of organic solvent nanofiltration (OSN) in order to narrow pore size distribution and offer high solvent permeance at low molecular weight cut offs. Functionalising the α-CD monomers with amine groups allowed for the primary end of the intrinsic pore to exclusively participate in the interfacial polymerisation reaction, resulting in the novel alignment of the intrinsic pores. Atomic force microscopy results indicate the selective thin films are ultra-thin, with thicknesses ranging from 6.5nm to 30nm. Molecular weight cut offs between 300-320 g mol^{-1} were achieved, demonstrating the applicability of the membranes for OSN. Permeance of methanol and ethanol were reasonable, ranging from 0.5 to 4 L m^{-2} hr^{-1} bar^{-1}, while permeance of n-hexane and toluene were much lower, in the range of 0.1 to 0.5 L m^{-2} hr^{-1} bar^{-1}.

1 Introduction

40-70% of the operating and capital costs of both the chemical and pharmaceutical industries is taken up by separation processes [1]. Not only do phase-induced chemical separations require large capital and operating costs, they are also very energy intensive [2]. In these industries, organic solvents can be found as both carriers of high value products, such as active pharmaceutical ingredients, and as raw feeds which are desired to be recycled.

Membrane separations offer an alternative to the costly and energy intensive phase induced separations, capable of performing the same separations while requiring an order of magnitude less energy [2]. Designing membranes that are stable in organic solvents has become an area of high interest in the field of membrane science, a use-case which has been deemed organic solvent nanofiltration (OSN) [3].

OSN membranes typically take the form of thin film composite (TFC) membranes, in which the selective and supporting layers are decoupled. The selective layer is ultra-thin, due to its formation via interfacial polymerisation, and determines membrane performance characteristics such as molecular transport and rejection. The support layer is porous and is used to enhance the mechanical stability of the selective layer.

The decoupling of the selective and supporting layers offers the ability to optimise each layer separately, where design criteria are: i) tunable pore size; ii) a narrow pore size distribution; iii) suitable interactions between permeants and the membranes; iv) an ultra thin selective layer (in the nanoscale) for enhanced permeance [4]. Efforts to achieve such a design either focus on utilising building blocks of intrinsic porosity, or preventing the efficient packing of the polymer chains. The former has been approached by using materials such as metal organic frameworks, covalent organic frameworks, and carbon nanosheets, while research using the latter approach makes use of technologies such as polymers of intrinsic porosity and carbon nano-spacers [5–9].

With these criteria in mind, this work proposes the use of α-cyclodextrin (α-CD) molecules as the building blocks of the selective layer in conjunction with terephtaloyl chloride (TPC) monomers. The result is the formation of an ultrathin selective layer network which possess hydrophilic and hydrophobic channels as well as intrinsic pores of fixed dimensions. The presence of both hydrophilic and hydrophobic channels in the membrane network has been deemed a *Janus Pathway* due to the opposing natures of the channels [4].

Cylcodextrin (CD) is an inexpensive glucose macrocycle that is nontoxic and may be made sustainably from starch [10, 11]. Its use as a building block in the selective layer of TFC membranes has been investigated following the trend of utilising macromolecules of intrinsic porosity, as these offer well defined pores helping to create fixed and uniform free volumes within the membrane network. The fixed intrinsic volume offers a pathway to obtaining narrow pore size distributions which increases the selectivity of separations, while the three variants of CD (α-, β-, and γ-) are of varying aperture sizes meaning the intrinsic pore size is tunable. This study focuses on the smallest of the three variants, α-CD, which has a pore aperture of 4.7-5.7Å[4].

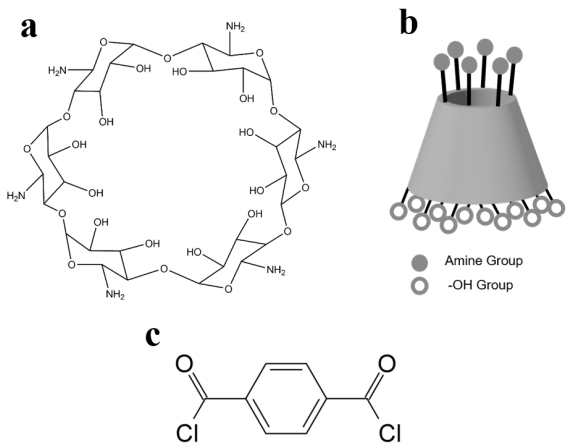


Figure 1: Monomer structures: **a**) Chemical structure of functionalised α-CD, **b**) 3-D schematic of functionalised α-CD, **c**) Chemical structure of TPC.

Since Villalobos et al. presented their work successfully synthesising CD – TPC membranes on a PAN support using an in-situ interfacial polymerisation method in 2017, there has been a large increase in the number of studies using CD as a building block in TFC membranes [4, 12–18]. While some of these studies report modifications to the CD molecule in order to maximise certain properties such as hydrophilicity, the CD molecules in these reported membranes are randomly packed [14]. A modification in attempt to align the CD molecules is yet to be reported, such is the novelty of this paper.

The primary hydroxyl groups of the α-CD molecule have been replaced with amine groups, the resulting modified molecule can be seen in Figures 1(**a, b**). During the interfacial polymerisation reaction with TPC, the amine groups seletively react with the TPC's acyl-chloride groups due to the neutral pH of the aqueous phase. This forms a membrane network with the functionalised (amine) ends facing the exterior selective layer surface and the hydroxyl ends facing the supporting membrane. This alignment has been confirmed by NMR in an unpublished study conducted by the Livingston Group at Imperial College London.

The hypothesis is that by aligning the pores of the CD molecules to be facing the surface, there would be an increase in the fraction of channels shown directly to the feed solution that would be hydrophobic, reducing a non-polar solvent pathway's tortuosity, which reduces resistance and increases flux. This would effectively increase the permeability of non-polar solvents due to the favorable interactions, as well as create tighter rejections due to the intrinsic porosity of the α-CD building block.

2 Methods

2.1 Materials

α-CD was purchased from Tokyo Chemical Industry UK Ltd. Solvents used were all HPLC grade, supplied from VWR international. Solvents used include: n-hexane; methanol; ethanol and toluene. TPC 98% was supplied from Sigma Aldrich. The PAN support membrane was made by the Livingston Group at Imperial College London.

Entry no.	Solution	Label	Composition % (w/w)
1	Organic (n-hexane)	1	0.1% Terephtaloyl chloride (TPC)
2	Aqueous (water)	2a	0.1% α-Cyclodextrin (α-CD)
3	Aqueous (water)	2b	0.2% α-Cyclodextrin (α-CD)
4	Aqueous (water)	2c	0.5% α-Cyclodextrin (α-CD)

Table 1: Solutions used for synthesis of TFC membranes

2.2 Preparation of solutions

Solution 2 is prepared by dissolving α-CD in DI-water (Solution 2a, 2b and 2c) and placing solution in a sonicator at room temperature until complete dissolution. Solution 1 was prepared utilising the same method. Solution 1 was made fresh on the same day as use to avoid hydrolysis which leads to a reduction in reactivity of the TPC molecules. 0.1% (w/w) TPC was chosen based on the solubility limit of TPC in n-hexane; 0.4 - 1% (w/w) TPC n-hexane solution did not exhibit complete dissolution of TPC while 0.05 - 0.4% (w/w) exhibited complete dissolution and observable membrane formation with α-CD. The latter was determined in conjunction with literature [4].

2.3 Preparation of α-CD – TPC membranes (interfacial polymerisation)

30ml of Solution 2(a, b or c) was added to a petri dish, enough solution to cover the metal plate used to transfer the film. Solution 1 was then slowly poured into the petri dish, forming an interface between the two immiscible solutions. The interfacial polymerisation reaction was allowed to occur for 1 minute. The film was transferred to a water bath using the immersed metal support layer in order to remove unreacted acyl-chloride groups from the surface of the film. A PAN membrane was taped to a plastic support and used to pick up the washed film in order to form the TFC membrane. The TFC membrane was then left to dry for 12 hours.

2.4 Membrane characterisation

2.4.1 Atomic force microscopy (AFM)

AFM provides a height measurement corresponding to the thickness of the α-CD–TPC selective layer. A smooth free-standing nanofilm formed via interfacial polymerisation is transferred onto a silicon wafer and scratched in order to determine the height profile of the selective layer. The AFM results were courtesy of the London Centre of Nanotechnology housed at University College London.

2.4.2 Water contact angle

KRÜSS sessile drop analysis was used to produce droplets and deduce the contact angle of water and the selective layer surface of the TFC membrane. Measurement was taken for 10 seconds at 10 fps. A mean of the angles tangent to the bubble surface at the membrane interface was then taken and the standard deviation of the 100 data points was then calculated.

2.4.3 UV/vis

Analysis and quantification of dye rejection was obtained using UV Probe software on a PC-controlled Shimadzu 1800 UV/vis system. The absorbance reading at the characteristic wavelength of adsorption of the dye is recorded for both the feed and permeate solutions, and the rejection is calculated using eq. 1.

$$R_i = (1 - \frac{Abs_{P,i}}{Abs_{F,i}}) \cdot 100\% \quad (1)$$

Where R corresponds to the rejection, given as a percentage; Abs corresponds to the absorbance of the permeate (Abs_P) and feed solutions (Abs_F) and subscript i represents the different dyes tested.

2.5 Membrane performance

Membrane performance is characterised according to two different criteria: (i) the permeance of four different solvents; two of which are polar (methanol and ethanol) and two non-polar (n-hexane and toluene) and; (ii) the molecular weight cut off

(MWCO) curve. Both permanence and rejection characteristics utilise the Sterlitech HP4750 Stirred Dead-end Cell applying a pressure of 10 bar to the surface of the feed liquid with the cell maintained at 25°C via a hot water bath.

2.6 Dye rejections

Before the dye rejection tests were carried out, the membrane was activated by feeding pure methanol solvent through the membrane for one hour. This serves two purposes: firstly, the membrane is compressed, stabilising the permeance as the free volume is reduced to a steady state; secondly, the free volumes are saturated with methanol solvent (the same solvent used to dissolve the dye molecules). Dyes tested and their corresponding molecular weights are displayed in Table 2.

Dye	Molecular Weight [g/mol]
Trolox	250
Methyl Orange	327
Sunset Yellow	452
Acid Fuchsin	585

Table 2: Molecular weights of dyes

When testing dye rejections, a stirrer was operated at 250rpm to prevent concentration polarisation at the feed-side surface of the membrane. Polarisation at the membrane surface would otherwise influence results producing lower permeance and possibly greater rejection. The permeate solution was tested under UV/vis where absorbance was used as an indicator for dye concentration. The MWCO corresponds to the molecular weight that exhibits a 90% rejection.

2.7 Solvent permeance

Solvent permeance results were calculated, with units L hr^{-1} m^{-2} bar^{-1}, following the dye rejection tests via the following equation:

$$Permeance = \frac{\dot{Q}}{P \cdot A_{cs}} \tag{2}$$

where $\dot{Q}$ represents the volumetric flowrate of permeate recorded in L h^{-1}; P is the pressure in bar and A_{cs} is the cross sectional area of the membrane in m^2 (fixed at the internal area of the dead-end cell at $1.256 \cdot 10^{-3}$ m^2). Each solvent was left to permeate for 1 hour before taking volumetric flowrate measurements to saturate the free volumes with the solvent of interest.

3 Results & Discussion

3.1 Membrane characterisation

Each α-CD–TPC membrane was characterised based on the hydrophobicity and thickness results obtained from water contact angle and AFM tests respectively. These in turn allowed conclusions to be drawn regarding the permeance of various polar and non-polar solvents. Water contact angles and average membrane thicknesses are displayed in Table 3 and Figure 2 respectively.

3.1.1 Selective layer thickness

The thickness of the α-CD selective layer of the membrane gives insight to the mass transfer mechanisms and context to the permeance results found. From thinnest to thickest, 0.5%, 0.1% and 0.2% (w/w) α-CD displayed thicknesses of 6.63nm, 15.28nm and 29.68nm respectively.

Thickness varies non-linearly with the concentration of α-CD in aqueous solution, suggesting that there are competing factors influencing the thickness of the membrane formed. Reaction duration is considered a key influence on the thickness of a membrane produced via interfacial polymerisation; a longer reaction time allows for more monomer to diffuse through the newly formed selective film to the reaction plane located at the interface between the two immiscible liquids [4, 12]. Since reaction time was controlled at a constant value of 1 minute, we can deduce this relationship is determined by other factors.

The rate of membrane formation is determined by competing factors: (i) the rate of diffusion to the reaction site and (ii) the kinetics of the nucleophilic acyl substitution reaction. As the activities of the monomers in the solution decreases, the rate of reaction becomes increasingly reliant on the mass

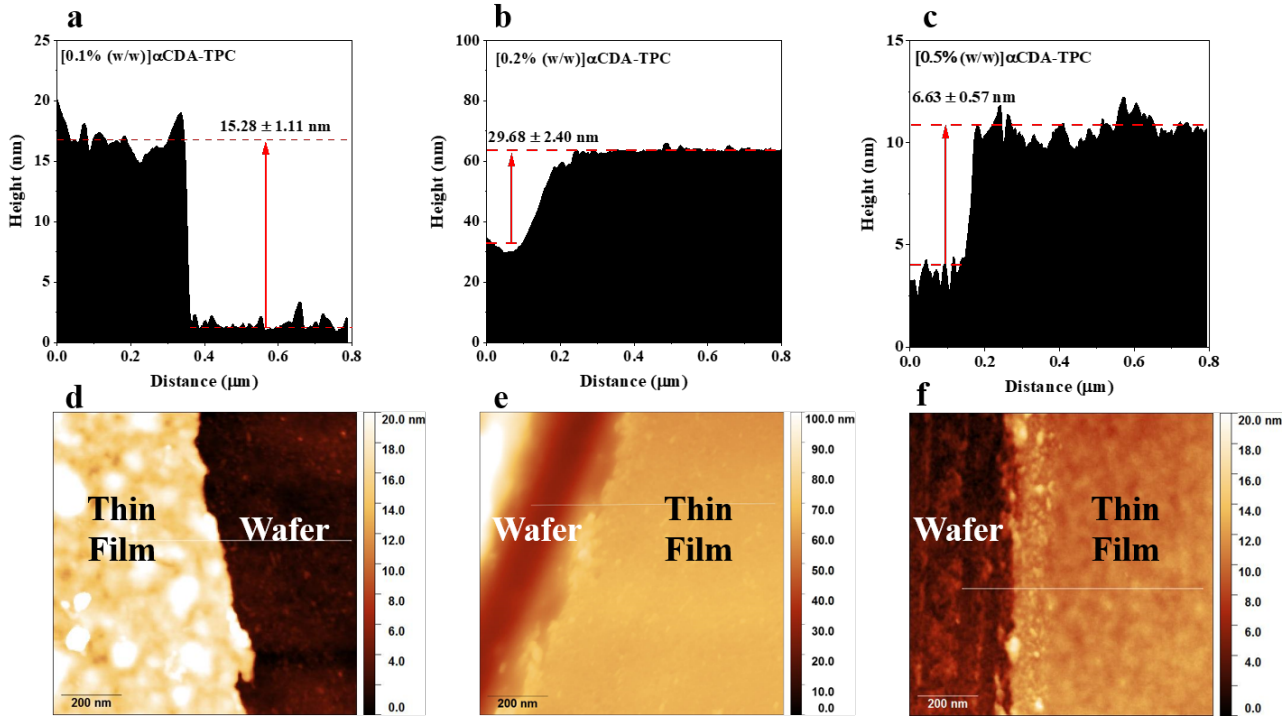


Figure 2: AFM results for the membranes of varying α-CD concentrations. In the top row: step change indicates the jump from the supporting silicon layer to the top of the selective film layer. Plot of measured thickness of (w/w) α-CD-TPC membrane concentrations of: **a**) 0.1%, **b**) 0.2%, **c**) 0.5%. In the bottom row: dark regions indicate the silicon wafer and the lighter ones indicate the selective film. AFM image for α-CD-TPC membrane concentrations of: **d**) 0.1%, **e**) 0.2%. **f**) 0.5%

transfer effects in the solution. Thickness is an indicator of the total monomer reacted during the interfacial polymerisation reaction. It then follows that thickness is expected to increase with concentration of monomer solution; unexpectedly, the results realised do not display increasing thickness with concentration of α-CD in Solution 2 [4]. Instead, 0.5% (w/w) α-CD was the thinnest membrane at 6.63nm, (8 α-CD layers), and 0.2% (w/w) was the thickest, reaching a height of 29.68nm, (38 α-CD layers).

At high concentrations of CD, agglomerations in aqueous solution become increasingly favourable [19]. Agglomerations cause activities of monomer in solutions to decrease since the reactant is effectively removed from solution. Furthermore, studies concerned with the relationship of CD diffusion and concentration in aqueous solution have noted decreases in activities and diffusion coefficients with increasing CD concentrations [20]. Increasing concentration will both increase the flux of α-CD towards the interface and increase the tendency for it to agglomerate. Hence, an optimal concentration of α-CD would be consistent with these competing effects.

NaOH can be used to break up agglomerations of CD and remove a proton from the hydroxly group to make CD more reactive [12]. Since NaOH was not used, as the functionalised α-CD has increased solubility in neutral pH, agglomerations of α-CD are more likely to group together and thus remove reactant from solution.

Further study is required to determine the concentrations of α-CD which provide the thinnest feasible film in order to produce thin selective layers of high permeability.

3.1.2 Surface hydrophobicity

To increase permeance of non-polar solvents, hydrophobic surfaces and free volumes are required [21]. Cyclodextrins offer a route to such a material through their intrinsic porosity, as seen in Figures 1(**a, b**). Being a macrocycle of glucose molecules, a CD

pore is hydrophobic, though the exterior network it forms with TPC is hydrophilic [22].

Membrane	Water Contact Angle
PAN	$37.6 \pm 1.28°$
0.1% (w/w) α-CD-TPC	$74.8 \pm 0.93°$
0.2% (w/w) α-CD-TPC	$40.5 \pm 0.62°$
0.5% (w/w) α-CD-TPC	$79.7 \pm 1.14°$

Table 3: Results of the water contact angle for PAN and the α-CD-TPC membranes of various concentrations

The water contact angle results presented in Table 3 describe a nonlinear increase in contact angle with an increase in α-CD concentration, where an increase in water contact angle indicates an increase in surface hydrophobicity [23]. Comparing the water contact angle of the PAN supporting membrane, 37.6°, to the contact angles of the 0.1%, 0.2% and 0.5% (w/w) α-CD films, 74.7°, 40.5°, and 79.7° respectively, it is apparent that the CD film increases surface hydrophobicity regardless of the concentration present. The 0.1% and 0.5% (w/w) membranes provide contact angles that are very similar to other reported free standing thin film CD membranes, the nonlinear increase in such contact angles is also corroborated by other reported studies though to a lesser extent [4, 15]

In a study investigating the hydrophobicity of TFC membranes for use in OSN, a TFC control membrane is reported to have a contact angle of 48.1° [21]. This contact angle is indicative of the hydrophilic nature of a conventional TFC membrane (TPC-MPD). Comparing this contact angle to those achieved during this study, it is apparent that the 0.1% and 0.5% (w/w) α-CD membranes achieved impressive hydrophobicity, while the 0.2% (w/w) α-CD membrane reports fairly strong hydrophilicity.

A possible explanation for the hydrophilic 0.2% (w/w) film lies in the thickness of the film reported in Figure 2(**a**). The 0.2% (w/w) film has a thickness twice as large as the other two membranes, indicating there are far more layers of α-CD–TPC than in the other two films. As thickness increases, it is expected that the alignment of the CDs is reduced as the probability of random packing increases upon each layer formed. If the α-CD molecules are randomly packed, there are likely to be fewer intrinsic pores facing the film surface, meaning there are fewer hydrophobic pathways presented to the water droplet. Surface roughness may also contribute to higher contact angles, however the large deviation in 0.2% (w/w) AFM results show that it is the roughest of the three samples tested [24].

While thickness is a possible explanation, comparison with the water contact angles of reported randomly packed CD membranes means that the effect this would have on water contact angle would be far less pronounced than a decrease of around 30°. To validate these results, it is necessary to further investigate this behavior by repeating these studies with more α-CD concentrations.

3.2 Membrane Performance

The performance of the various membranes were characterised by permeance (measure of solvent transport) and rejection (measure of ability to separate solutes from solvent). The results of both studies are presented in Figure 3.

3.2.1 Solute rejection

Solute rejection was conducted using molecular weight was used as an indicator of molecule size. A molecular weight cut off (MWCO) curve was constructed for the various membranes, where a molecule is considered rejected if separation is >90% (the cut off, indicated by a dashed line in Figures 3(**a, b, c**)). A sharp cut off is desirable as it allows for precise and selective separations of solutes on either side of this cut off weight [3]. The sharpness of a cut of curve is indicative of the steepness of the slope in Figures 3(**a, b, c**). α-CD concentrations of: 0.5% (w/w) displayed the steepest cut off; 0.1% (w/w) displayed the shallowest while 0.2% (w/w) displayed an intermediate cut off. These properties align with the hypothesis that the thinner the selective layer is, the more aligned the α-CD molecules are. Increased alignment, in turn, results in a membrane with increased selectivity to molecules smaller than the intrinsic α-CD pore diameter (4.7Å).

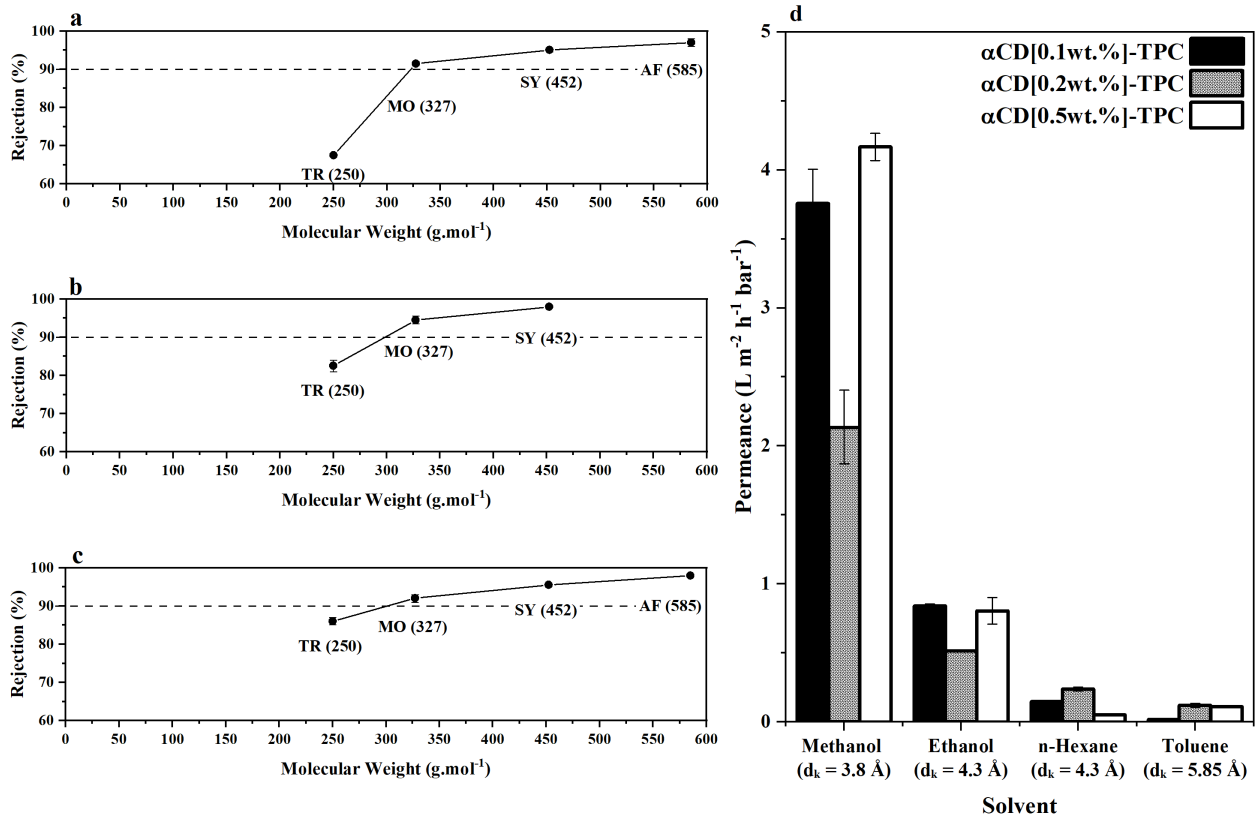


Figure 3: Molecular weight cut off (MWCO) curves of the TFC membranes of varying α-CD concentrations using four dyes dissolved in methanol at 20ppm: trolox (TR), methyl orange (MO), sunset yellow (SY), and acid fuchsin (AF). MWCO curve for (w/w) α-CD concentrations of: **a**) 0.1% MWCO of 320 g.mol^{-1}, **b**) 0.2% MWCO of 300 g.mol^{-1}, **c**) 0.5% MWCO of 300 g.mol^{-1}. **d**) Nanofiltration performance of the TFC membranes of varying α-CD concentrations in methanol, ethanol, n-hexane and toluene, with kinetic diameters of each solvent listed [25–27].

0.5% and 0.1% (w/w) α-CD display very similar rejections for dyes above the MWCO (92% and 91%, 95.5% and 95%, 98% and 97% for methyl orange, sunset yellow, acid fuchsin respectively) presenting evidence for the alignment of α-CD resulting and a well defined pore size distribution of the membrane. The thickest, and suspected least aligned membrane, 0.2% (w/w) α-CD, displays rejections less alike the other two (with 94.5% and 98% for methyl orange and sunset yellow). Below the MWCO, higher rejections are achieved with increasing thickness as rejection becomes increasingly dependent on thickness rather than pore size. A thicker membrane provides larger resistance to dyes that can permeate through the membrane, therefore smaller molecular weight dyes are able to permeate through less. [28]. These results reiterate the need for optimisation of α-CD concentration to achieve ultra-thin membranes of a narrow pore size distribution.

3.2.2 Polar and non-polar solvent permeance

Transport through a membrane can be described by a three step process: i) a molecule dissolves into the feed side of the membrane, ii) the molecule diffuses through the membrane's free volume, iii) the molecule diffuses out of the membrane into the the permeate [29]. Hence, the hydrophilicity and hydrophobicity of a membrane is key in determining whether a membrane is appropriate for use for filtration of polar and non-polar organic solvents respectively.

Polar solvents experienced the highest permeate flux through the various membranes, as displayed in Figure 3(**d**). In order of increasing selective layer thickness, the permeances of methanol were: 4.2, 3.8 and 2.1 L m^{-2} hr^{-1} bar^{-1}; and ethanol: 0.81, 0.84

and 0.51 L m^{-2} hr^{-1} bar^{-1}, for 0.5%, 0.1% and 0.2% respectively. Decreasing selective layer thickness is expected to increase the permeance through the membrane. The results for methanol are as expected, as are those for ethanol, within the reasonable error presented in Figure 3(**d**). The need to fully saturate a membrane's free volumes with the permeate of interest is a potential explanation for the notable error in ethanol permeance of 0.5% (w/w) α-CD membranes. Each polar solvent is dissolved at the hydrophilic surface site and diffuses through the pathways formed between the TPC molecules and the exterior of the α-CD molecules. Decreased permeance of ethanol relative to methanol is resultant from its larger kinetic diameter.

The opposite relationship was found for toluene, with the thickest membrane, 0.2% (w/w) α-CD, displaying the largest toluene permeance of 0.12 L m^{-2} hr^{-1} bar^{-1} compared to 0.11 and 0.02 L m^{-2} hr^{-1} bar^{-1} for 0.5% and 0.1% respectively. Toluene has a kinetic diameter of 5.85Å; this value is larger than the diameter of the hydrophobic pathway, that is the internal diameter of α-CD of 4.7Å[27]. Hence, the permeance of toluene is as expected if the hydrophobic species were to pass through the hydrophobic inner cavity of the aligned α-CD. Toluene is too large to interact with the hydrophobic internal cavity of α-CD and therefore is unable to permeate effectively in the case of the highly aligned 0.5% and 0.1% (w/w) α-CD selective layers. The increased permeance of toluene through the thinnest aligned α-CD, 0.5% (w/w) membrane is put down to the reduced resistance of the thinner selective layer compared to the thicker, 0.1% (w/w) α-CD aligned selective layer. While it is possible that the random packing of the 0.2% (w/w) α-CD thin film meant there was a broader pore size distribution due to the variance of the α-CD–TPC network, meaning there were more pathways able to facilitate the transport of toluene molecules.

n-Hexane was chosen for direct comparison to ethanol, with the aim of comparing the permeance for a polar and non-polar solvent of the same kinetic diameter. As displayed in Figure 3, the two membranes considered to be the most highly aligned, 0.1% and 0.5% (w/w) α-CD, experienced a reduction in permeance for n-hexane compared to ethanol. This implies one or both of: (i) a large ratio of hydrophilic to hydrophobic free volume or (ii) a wide pore size distribution for hydrophilic relative to hydrophobic pathways. The TPC monomer making up the hydrophilic pathways have no intrinsic porosity, hence the pore size distribution is influenced by the mechanism of the selective layer formation. From this, the larger permeance experienced by ethanol is hypothesised to be resultant from the wide distribution of hydrophilic pore sizes relative to the hydrophobic inner cavity pathway.

This warrants further investigations: the effect of increasing the size of the CD intrinsic pore on permeation of non-polar organic solvents; and quantifying the pore size distribution of the functionalised α-CD membranes via techniques such as BET. For the non-polar OSN, exchanging the CD monomer for one with a larger inner cavity radius could dramatically increase the flux of the solvent through the selective layer, and using a monomer with inherent porosity smaller than the kinetic diameter of the solvent could produce sharp and precise cut offs for larger non-polar organic solvents.

4 Outlook

During this study, free standing non-functionalised (pristine) α-CD membranes were made via interfacial polymerisation for the sake of comparison to functionalised α-CD membranes. However, testing performance of the pristine membranes presented issues regarding their mechanical stability that were not encountered with functionalised α-CD membranes.

Repeating dye rejections resulted in dramatic reductions in rejection performance upon subsequent tests. For example, pristine α-CD sunset yellow rejections fell from 97% to 93% and finally to 25% in subsequent tests.

This warrants investigation into the mechanical stability of functionalised CD membranes and pristine alternatives. Mechanical stability and the ability for membranes to maintain constant performance over

their life-cycle is extremely important for industrial membranes. Further investigation into the potential stability advantages of alternative CD membranes is also relevant alongside the comparison of performance characteristics to draw conclusions on the superiority of each alternative.

5 Conclusions

This work demonstrates the potential for amine-functionalised α-CD–TPC membranes for use in OSN. The novel alignment of the intrinsic pores is a promising route to obtaining larger degrees of hydrophobicity; narrower pore distributions; and decreased overall resistances to solvent transport in selective layers of TFC membranes. The ultra thin films (<30nm) presented in this work demonstrate the advantages of using molecules of intrinsic porosity, as these thin films were still able to successfully separate small molecules, with MWCO's around 300 g mol^{-1}. Permeance of both polar and non-polar solvents was found to be adequate for membranes with MWCO's around 300 g mol^{-1}.

This work targets a very new and innovative area of membrane science with large potential. Lack of previous work on amine-functionalised CD membranes results in the requirement for further investigation to characterise and fully understand the results presented in this report.

6 Acknowledgements

The authors are grateful to the members of the Livingston Group at Imperial College London for the guidance and support provided during the project.

References

[1] S. Adler, E. Beaver, P. Bryan, S. Robinson, and J. Watson. Vision 2020:2000 separations roadmap. *Center for Waste Reduction Technologies of the AIChE and Dept. of Energy of the U.S.A*, 2000.

[2] D.L. Gin and R.D. Noble. Designing the next generation of chemical separation membranes. *Science*, 332(6030):674–676, 2011.

[3] P. Marchetti, M.F. Jimenez Solomon, G. Szekely, and A.G. Livingston. Molecular separation with organic solvent nanofiltration; a critical review. *Chemical Reviews*, 114(21):10735–10806, 2004.

[4] J. Liu, D. Hua, Y. Zhang, S. Japip, and T-S. Chung. Precise molecular sieving architectures with janus pathways for both polar and nonpolar molecules. *Advanced Materials*, 30(1705933), 2018.

[5] M.S. Denny Jr, J.C. Moreton, l. Benz, and S.M. Cohen. Metal-organic frameworks for membrane-based separations. *Nature Reviews Materials*, 1(16078), 2016.

[6] M.S. Denny Jr, J.C. Moreton, l. Benz, and S.M. Cohen. Covalently linked metal-organic framework (mof)-polymer all-solid-state electrolyte membranes for room temperature high performance lithium batteries. *Journals of Material Chemistry A.*, 6:17227–17234, 2018.

[7] M.F. Jimenez Solomon, Q. Song, K.E. Jelfs, M. Munoz Ibanez, and A.G. Livingston. Polymer nanofilms with enhanced microporosity by interfacial polymerization. *Nature Materials*, 15: 760–767, 2016.

[8] S. Karan, S. Samitsu, X. Peng, K. Kurashima, and I. Ichinose. Ultrafast viscous permeation of organic solvents through diamond-like carbon nanosheets. *Science*, 335:444–447, 2012.

[9] M. Cook, P.R.J. Gaffney, L.G. Peeva, and A.G. Livingston. Roll-to-roll dip coating of three different pims for organic solvent nanofiltration. *Journal of Membrane Science*, 558:52–63, 2018.

[10] E.M. Martin Del Valle. Cyclodextrins and their uses: a review. *Science*, 39(9):1033–1046, 2004.

[11] H. Dodziuk. *Cyclodextrins and Their Complexes: Chemistry, Analytical Methods, Applications*. Wiley-VCH Verlag GmbH & Co., Weinheim, Germany, 2006.

[12] LF. Villalobos, T. Huang, and K-V. Peinemann. Cyclodextrin films with fast solvent transport and shape-selective permeability. *Advanced Materials*, 29(1606641), 2017.

[13] S. Xiong, D.Y. Zhang, S. Mei, J. Liu, Y.S. Shi,

and Y. Wang. Thin film composite membranes containing intrinsic cd cavities in the selective layer. *Journal of Membrane Science*, 551:294–304, 2018.

[14] S. Xiong, S. Xu, S. Zhang, A. Phommachanh, and Y. Wang. Highly permeable and antifouling tfc fo membrane prepared with cd-eda monomer for protein enrichment. *Journal of Membrane Science*, 572:281–290, 2019.

[15] J. Xue, Z. Jiao, R. Bi, R. Zhang, X. You, F. Wang, L. Zhou, Y. Su, and Z. Jiang. Chlorine-resistant polyester thin film composite nanofiltration membranes prepared with β-cyclodextrin. *Journal of Membrane Science*, 584:282–289, 2019.

[16] Y. Zhang, H. Sun, H. Sadam, Y. Liu, and L. Shao. Supramolecular chemistry assisted construction of ultra-stable solvent-resistant membranes for angstrom-sized molecular separation. *Chemical Engineering Journal*, 371:535–543, 2019.

[17] X. Wu, Y. Chen, and Y. Liu. Supramolecular crosslinked polymer for efficient organic dye removal from aqueous solution. *Advanced Sustainable Systems*, 3(1800165), 2019.

[18] S-J. Xu, Q. Shen, G-E. Chen, and Z-L. Xu. Novel β-cd@zif-8 nanoparticles-doped poly(*m*-phenylene isopthalamide) (pmia) thin-film nanocomposite (tfn) membrane for organic solvent nanofiltration (osn). *ACS Omega*, 3: 11770–11787, 2018.

[19] M. Bonini, S. Rossi, G. Karlsson, M. Almgren, P. Nostro, and P. Baglioni. Self-assembly of â-cyclodextrin in water. part 1: Cryo-tem and dynamic and static light scattering. *Langmuir*, 22:1478–1484, 2005.

[20] A.C.F. Ribeiro, Valente A.J.M., C.I.A.V. Santos, P.M.R.A. Prazeres, V.M.M. Lobo, H.D. Burrows, M.A. Esteso, Cabral A.M.T.D.P.V., and F.J.B. Veiga. Binary mutual diffusion coefficients of aqueous solutions of α-cyclodextrin, 2-hydroxypropyl-α-cyclodextrin, and 2-hydroxypropyl-β-cyclodextrin at temperatures from (298.15 to 312.15) k. *Journal of Chemical Engineering Data*, 52(2):586–590, 2007.

[21] M.F. Jimenez Solomon, Y. Bhole, and A.G. Livingston. High flux hydrophobic membranes for organic solvent nanofiltration (osn)-interfacial polymerization, surface modification and solvent activation. *Journal of Membrane Science*, 434(193-203), 2013.

[22] S. Xiong, S. Xu, S. Zhang, A. Phommachanh, and Y. Wang. Highly permeable and antifouling tfc fo membrane prepared with cd-eda monomer for protein enrichment. *Journal of Membrane Science*, 572:281–290, 2019.

[23] S. Wang, J. Cai, W. Ding, Z. Xu, and Z. Wang. Bio-inspired aquaporinz containing double-skinned forward osmosis membrane syntehsized through layer-by-layer assembly. *Membranes*, 5:369–384, 2015.

[24] B. Bhushan and Jung Y.C. Micro- and nanoscale characterization of hydrophobic and hydrophilic leaf surfaces. *Nanotechnology*, 17(11), 2006.

[25] B. Bettens, S. Dekeyzer, B. Van der Bruggen, J. Degreve, and C. Vandecasteele. Transport of pure components in pervaporation through a microporous silica membrane. *Journal of Physical Chemistry*, B:5216–5222, 2005.

[26] A.F.P. Ferreira, M.C. Mittelmeijer-Hazeleger, M.A. Granato, V.F. Duarte Martins, A.E. Rodrigues, and G. Rothenberg. Sieving di-branched from mono-branched and linear alkanes using zif-8: experimental proof and theoretical explanation. *Physcial Chemistry*, 15: 8795–8804, 2013.

[27] Y. Weng, S. Qiu, L. Ma, Q. Liu, M. Ding, Q. Zhang, Q. Zhang, and T. Wang. Jet-fuel range hydrocarbons from biomass-derived sorbitol over ni-hzasm-5/sba-15 catalyst. *Catalysts*, 5 (2147-2160), 2015.

[28] G.M. G, H.B. Park, A.C. Saglea, B.D. Freeman, and J.E. McGrath. Water permeability and water/salt selectivity tradeoff in polymers for desalination. *Journal of Membrane Science*, page 130–138, 2011.

[29] H.B. Park, J. Kamcev, L.M. Robeson, M. Elimelech, and B.D. Freeman. Maximizing the right stuff: The trade-off between membrane permeability and selectivity. *Science*, 5(6343), 2017.

Techno-Economic Assessment of the Direct Air Capture of Carbon Dioxide

Su Cheng Tung and **Jamie MacNeil**

Department of Chemical Engineering, Imperial College London, UK

December 2019

Abstract

CO_2 Direct Air Capture has been long investigated as a suitable candidate in the options for climate change mitigation. However, with cost being one of the largest barriers, the uptake of this technology has been limited. In this study, the first ever techno-economic analysis using both literature review and process modelling of a large-scale, Low Temperature Direct Air Capture plant is performed using ASPEN Plus. This process captures 10,000 t_{CO_2}/year from the atmosphere using a polyethylenimine adsorbent in a batch, fixed bed adsorption column. The technical design and economic analysis of each process unit is outlined, followed by possible downstream utilisations of CO_2 along with their associated costs. For this process, a capital expenditure of 930 $£{\cdot}year/t_{CO_2}$ was obtained, with a Levelised Cost of Direct Air Capture of 2782 £/t_{CO_2}. To predict how this figure may improve, a sensitivity analysis was performed. It was found that the cost of adsorbent contributed 90% of the total value. Therefore, with further improvements to the adsorbent and its production in bulk, it is believed that the Levelised Cost of Direct Air Capture can realistically be expected to fall to as low as 293 £/t_{CO_2}, encouraging wide uptake of Low Temperature Direct Air Capture technology.

1 Introduction

The average temperature of the Earth is rising faster than it has done for the past 2,000 years and the impacts include rising sea levels, more extreme weather and an increased rate of extinction. As a result, 195 countries signed the Paris Agreement in 2016 which set national and international goals to help prevent global warming. The main ambition of the Paris Agreement is to keep the global average temperature rise to less than 2°C above industrial levels and pursue further efforts to limit the increase to 1.5°C.

Greenhouse gases are primarily responsible for global warming and, of these, CO_2 has the largest anthropogenic impact. The average concentration of CO_2 in 2016 was found to have risen by 40% since the pre-industrial era and has been increasing by 2ppm/year for the past ten years (IEA 2017). Energy generation represents the largest source of CO_2 emissions and, since demand for energy comes from global economic growth and development, this will inevitably result in an upward trend in CO_2 emissions in the years to come (IEA 2018). In order to meet the Paris Agreement, global energy related emissions must be halved by 2050, as concluded by IEA (2012). Therefore, significant decarbonisation is essential.

There are several options available for CO_2 emission reduction, one of them being point source carbon capture. However, in addition to the fact that some plants are too old to be retrofitted with point source carbon capture, the average carbon capture efficiency ranges from 50-94%, so a significant amount of CO_2 is still released into the atmosphere. Another technique used for decarbonisation is fuel substitution. This successfully reduces emissions in the energy sector but emissions in the transport and industrial sectors remain significant (Leeson et al. 2017a). Although these are just a few examples of many technologies being employed to decarbonise, it highlights why they alone cannot completely mitigate anthropogenic CO_2 emissions.

DAC is a technology that removes CO_2 directly from the atmosphere. Using two Integrated Assessment Models (IAMs), Realmonte et al. (2019) estimated that the cost per t_{CO_2} required to meet the 1.5°C Paris Agreement could be reduced by between 66 and 98% with wide-scale deployment of DAC. Therefore, DAC should be explored as one of the long-term solutions for decarbonisation if the world is going to meet the Paris Agreement by 2050.

In this paper, a techno-economic assessment of a large-scale (10,000 t_{CO_2}/year) LT DAC plant coupled with modelling of process equipments is performed. The sections are laid out as follows: Section 2 discusses the past work in this field. Section 3 describes the methodology. Section 4 contains the results and discussion achieved in this study, with sub-sections for each process unit. Tables and figures are used to better illustrate these findings. After obtaining a final Levelised Cost of Direct Air Capture (LCOD), it is then contrasted to literature to see how well they compared. A sensitivity analysis is also carried out on the final LCOD varying the adsorbent's price and properties. Finally, downstream utilisations are discussed in section 5 and the conclusion is given in section 6, containing a summary of this study and further efforts to be made.

2 Background

CO_2 capture from air began in 1930 and was used for spaceships and submarines (Kurt Zenz House 2011). In 2006, David W. Keith (2005) described two DAC systems that could achieve a final LCOD of under 200 and 500\$/t_{CO_2}, Bioenergy Carbon Capture and Storage (BECCS) and aqueous alkaline solution-based DAC respectively. It was stated that the latter would be more easily implemented in the near future due the use of current technologies. However, the high heat demand (900°C) for aqueous alkaline solution-based DAC, referred to as High Temperature DAC (HT DAC) from now on, may limit its uptake due to cost and available heat sources. Significant research has been carried out in trying to optimise HT DAC,

including improvements to the current process, using alternative electricity sources, and even new process ideas with different numbers of cycles or alkaline solution used (Mahdi Fasihi 2019). Thanks to Maryam Mahmoudkhani (2009), a novel technique was discovered which managed to reduce the temperature required for HT DAC by at least 50°C, resulting in a substantial reduction in cost. In 2018, W.Keith et al. (2018) provided the first ever detailed techno-economic analysis of a real pilot plant which managed to capture 1Mt$_{CO_2}$/a.

Meanwhile, significant effort was also put into the research of a different type of DAC, Low Temperature (LT) DAC. As described by Sinha et al. (2017), Temperature Swing Adsorption (TSA) is the most common process used in LT DAC. By using solid sorbents to perform adsorption, relatively lower temperatures (80-100°C) are required to perform regeneration allowing for the use of free waste heat from power plants or processes running exothermic reactions, resulting in a reduction in cost. Depending on the adsorbent used, water can be produced as a side product in LT DAC, relieving the water demand constraint. Another method in LT DAC is Moisture Swing Adsorption (MSA) studied by Lackner (2009), where regeneration is carried out by moisturising the saturated adsorbents.

Other cutting-edge techniques suggested for DAC include electrochemical capture, the use of molecular filters, crystallisation and so on. However, these techniques only provide positive preliminary results, hence more research is required to consolidate their findings (Mahdi Fasihi 2019). There are also papers which discuss the available DAC technologies. Simon et al. (2011) analysed the LCOD of a typical DAC plant designed to capture 400t$_{CO_2}$/year using different energy sources. Micah Broehm (2015) made a comparison of 3 different DAC technologies (aqueous solutions, adsorbents and inorganic solid sorbents) and analysed their required consumables. Rather than comparing different DAC techniques amongst each other, Williamson (2016) discussed the difference between DAC and other CO_2-removal techniques.

Thanks to the comparison performed by Mahdi Fasihi (2019) of the predictions of LCOD, LT DAC was chosen as the focus of this study as it showed consistently lower values up to 2050. As shown by the literature review, existing research only considers partial techno-economic studies or, when a detailed analysis was available, it either referred to small-scale LT DAC plants or to HT DAC. Therefore, this is the first paper to design a process model, coupled with a techno-economic assessment, of a large scale LT DAC plant, by using data from both literature research as well as process modelling.

3 Methodology

The methodology presents a techno-economic study of the LT DAC process. First, the process design will be discussed. Then, the deciding factors on the adsorbent will be outlined. This is followed by the technical design of each process unit. Finally, the methodology for the economic analysis will be presented.

3.1 Process Flowsheet

A simplified process flowsheet is illustrated in figure 3.1. Air is removed from the atmosphere using extraction fans and sent to the dehumidifiers where its relative humidity is controlled. Solenoid valves are located after the dehumidifiers to prevent the inlet flow of air into the adsorption columns during desorption and three-way valves are used at the column exit to control the release of CO_2 depleted air back to the atmosphere (during adsorption) or transport the purified CO_2 downstream (during desorption). There is also a vacuum pump at the column exit to reduce the pressure in the column and extract the CO_2 during desorption. Finally, cooling water and waste heat streams are used to control the operating temperature of the adsorption column.

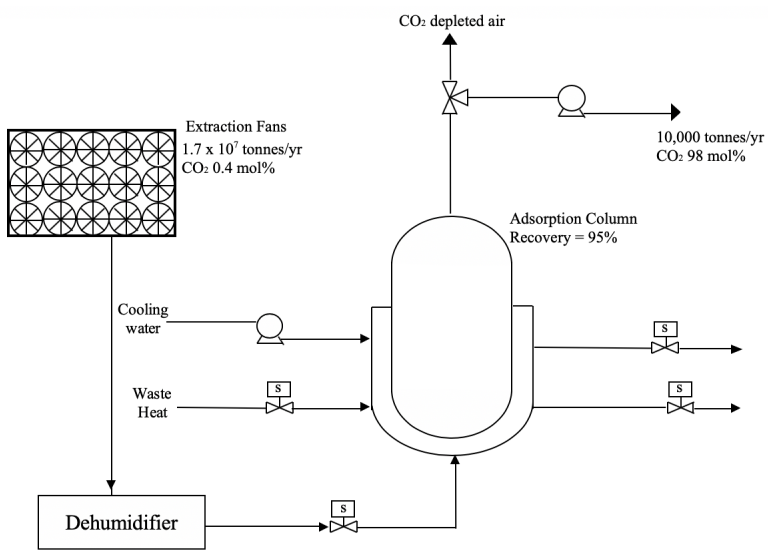


Figure 3.1: *DAC LCOD Predictions*

This process was designed to capture 10,000 t_{CO_2}/year from the atmosphere as higher capture rates would have resulted in excessive equipment sizing. Additionally, an outlet CO_2 purity of 98mol% was chosen in accordance with the findings by B. Wetenhall (2014). This paper showed that higher purities resulted in larger energy and column costs in the process, whereas lower purities resulted in larger compression costs and corrosion rates downstream of the process. A CO_2 recovery rate of 95% was also targeted.

3.2 Adsorbent Type

The adsorbent suitable for this study had to be able to capture CO_2 in concentrations as low as 400ppm. In this search, the main factors were adsorbent capacities, unit cost as well as degradation rate. The environmental impacts of the adsorbents were also considered. Since the adsorbents in this study are specifically functionalised to capture CO_2, they were not readily produced. The ratio of raw materials used to synthesise them along with their prices were used to calculate the unit cost of adsorbents. All Prices were taken from Sigma Aldrich.

3.3 Adsorption Column

Due to limited information on the adsorbent's properties, a rather simplified approach had to be taken when designing the adsorption column. The potential improvements of this methodology and their requirements are outlined in Section 6.2.

When estimating the size of the columns, a voidage of 0.37 was selected to allow for expansion of the adsorbent. Additionally, a cycle time of 13 hours was used, analogous to the experiment carried out by Wang et al. (2015). This included time for adsorption, desorption and heating. The density of the adsorbent was estimated using the weight fractions of its components and was found to be 888 kg/m^3. The column operated at atmospheric pressure given the lack of information available on how pressure affected saturation capacity. To find the most economically feasible design, the number of columns, height to diameter ratio of each column and replacement rate of the adsorbent was investigated using the following equations (symbols are defined in the nomenclature). Note that one batch was taken as the number of cycles before the adsorbent was replaced.

$$cycles\ per\ batch = 1 + \log_{1-dr} f \tag{1}$$

$$m_{CO_2,batch} = \frac{\dot{m}_{CO_2}}{batches\ per\ year} \tag{2}$$

A geometric series was used in the equation below to calculate how much adsorbent remained active after each cycle:

$$Cap_n = Cap_1 \times (1-d_r)^n \tag{3}$$

$$m_{CO_2,ads} = recovery\ rate \times \sum_{n=1}^{N} Cap_n \tag{4}$$

$$m_{ads,batch} = \frac{m_{CO_2,batch}}{m_{CO_2,ads}} \tag{5}$$

$$NC \times V = \frac{m_{ads,batch}}{\rho \times (1-\varepsilon)} \tag{6}$$

$$Cap.\ cost = 583.6 \times D^{0.675} \times H \times F_{mat} \times (\frac{P \times 14.5}{50})^{0.44} \tag{7}$$

Equation 7 is a correlation taken from E.Brunazzi (2002). The capital cost was given in euros (2002) but converted to pounds (2019) in accordance with the rest of the economic analysis. Because the column only operated at atmospheric pressure, the cheapest material (carbon steel) was used, which had a material correction factor (F_{mat}) of 1.

Equations 8 and 9 show how the outlet purity of the process was checked to ensure it was close to the target of 98mol%:

$$m_{air,col} = \varepsilon \times V \times \rho \tag{8}$$

$$CO_2\ Purity = 100\% \times \frac{C(CO_2)}{C(CO_2) + m_{air,col}} \tag{9}$$

3.4 Pumps and Heat Exchangers

ASPEN Plus was used to estimate the cost of heat exchangers and pumps. As the heat capacity of the adsorbent was not available, the short cut method was used. The total moles of CO_2 captured in a cycle was matched to the heat exchanger inlet so that adsorbent regeneration could be completed within an hour. A sensible jacket surface area was estimated using a percentage of the total surface area of the column. From there, the waste heat and cooling water flowrates were adjusted to meet the heat transfer area of the jacket. These flowrates were then used to estimate the cost of the cooling water pump. The costs associated with the vacuum pump placed at the product outlet was obtained from an online vendor (Kurt K. Lesker Company n.d.), while operational and utility costs were calculated using ASPEN Plus' price of electricity.

3.5 Valves

To achieve the objective of capturing 10,000 t_{CO_2}/year, the inlet flowrate of air was found. With the aim of achieving a reasonable Reynold's number, the quantity and area of inlet pipes required was calculated. The price of valves was found from (Flomatic Valves Company n.d.). An extrapolation was carried out to estimate the price of valves with diameter larger than 12'.

3.6 Dehumidifiers

Wang et al. (2015) investigated the saturation capacity of the adsorbent at various humidities. They found that the optimal relative humidity was 10%. Given that increasing the relative humidity to 60% reduced the adsorbent's capacity by roughly 25%, it was decided that the relative humidity of the inlet air should be rigorously controlled to 10%. This was achieved by adding dehumidifiers to the inlet air stream.

The number of dehumidifiers required was estimated using the inlet flowrate of the process and the maximum air flowrate that could be processed by each dehumidifier. Additionally, an extrapolation was carried out to find the capital costs and electrical requirements of large dehumidifiers (air flowrate greater than 50,000m^3/hr) using the available information for smaller dehumidifiers.

3.7 Extraction Fans

Heidel et al. (2011) investigated the costs of extraction fans for air capture. It was used to estimate their cost per tonne of CO_2 for this study. Their findings are presented in table 3.1.

Table 3.1: *Extraction Fan Methodology*

Cost Description	Value (£/t_{CO_2})
Capital Cost	8.63
Operational Cost	3.92
Electrical Cost	6.22

3.8 Economic Analysis

The methods used to cost each process unit are outlined in table 3.2. These were chosen based upon their accuracy and applicability.

Table 3.2: *Costing Methodology*

Unit	Costing Methodology
Adsorption Column	E.Brunazzi (2002)
Extraction Fans	Heidel et al. (2011)
Heat Exchangers	ASPEN Plus
Pumps	ASPEN Plus
Dehumidifiers	Online vendor
Valves	Online vendor
Vauum Pump	Online vendor

Following this, some assumptions had to be made. Firstly, the maintenance cost was taken as 5% of the capital cost because Leeson et al. (2017b) found that the maintenance cost of a plant is typically between 3 and 6% of the capital cost. However, this assumption was not used for the extraction fans, which had a more accurate estimate for their maintenance cost in Heidel et al. (2011). Using the findings of Mahdi Fasihi (2019), the Weighted Average Capital Cost (WACC) was taken as 7% and the plant lifetime was assumed to be 20 years. Finally, the Cost of Electricity (COE) was taken as 7.5 p/kWh to match the costing methods as used in Aspen Plus. These assumptions were used to calculate the annualised capital cost, capex and the LCOD:

$$crf = \frac{WACC \cdot (1+WACC)^l}{(1+WACC)^l - 1} \tag{10}$$

$$ACC = capital\ cost \cdot crf \tag{11}$$

$$Capex = \frac{capital\ cost}{\dot{CO_2}} \tag{12}$$

$$LCOD = \frac{ACC + OC + electrical\ input \cdot COE}{\dot{CO_2}} \tag{13}$$

4 Results and Discussion

4.1 Adsorbent Type

As shown in table 4.1, HP2MGL-CTAB was chosen due to its relatively high adsorption capacity and low rate of degradation, which resulted in fewer batches per year and finally the lowest total yearly cost. The environmental impacts of each adsorbents was also investigated, however they all had a negligible impact.

Table 4.1: *Adsorbents considered in this study*

Adsorbent type	Adsorption capacity	Degrad−ation	Batches	Yearly Cost
	mol_{CO_2}/kg_{ad}	%/cycle	-	mil.£/a
T-PEI[a]	2.19	0.61	3	33.8
A-PEI[b]	2.26	3.19	16	169
HP2MGL −CTAB[c]	**2.63**	**0.4**	**2**	**25.1**

a: Tetrapropyl Orthotitanate; b: 3-Aminopropyl Trimethoxysilane; c: Cetyl Trimethyl Ammonium Bromide

4.2 Adsorption Column

Both fixed bed (replacing all the adsorbent after a specified number of cycles) and fluidised bed columns (with a constant replacement rate and a recycle loop for the adsorbent) were investigated. Given the already large column volumes required, a fixed bed column was chosen due to its lower void fraction.

Figure 4.1 shows the results when replacing the adsorbent after varying degrees of degradation. It can be seen that replacing the adsorbent after more degradation results in a lower cost of adsorbent but a larger column volume. Therefore, it was decided to replace the adsorbent in the column when 30% remained active. This was a good compromise between adsorbent wastage, column size and adsorbent costs.

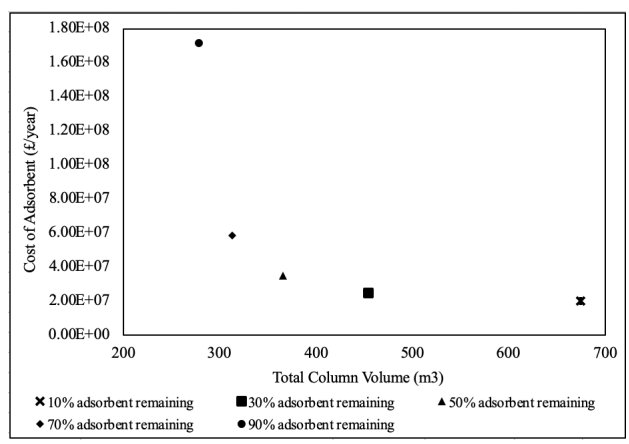


Figure 4.1: *Effect of Adsorbent Replacement Rate on the Total Column Volume and Cost of Adsorbent*

Figure 4.2 shows how the capital cost varies with the number of columns in the process and their height to diameter

(H/D) ratio. Purely from a cost perspective, the process would have been designed for 2 columns with a H/D ratio of 2 however, this was infeasible due to the excessive size of these columns (5.25 m in diameter and 10.5 m in height). Additionally, it should be noted that the ratio of the breakthrough time to saturation time of CO_2 decreased for lower H/D ratios, leading to a lower recovery rate of CO_2. Therefore, this process was designed with 6 columns and a H/D ratio of 5. These columns had a more manageable size (2.68 m in diameter and 13.4 m in height) and a higher recovery of CO_2 (95%). They also achieved an outlet purity of 98.8mol% of carbon dioxide which was only marginally larger than the optimal purity (98mol%) targeted in section 3.1. A summary of the column parameters can be seen in table 4.2.

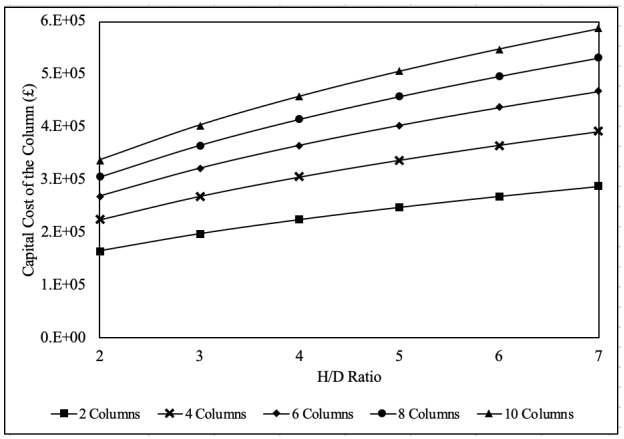


Figure 4.2: *Effect of the H/D Ratio and Number of Adsorption Columns on Capital Cost*

Table 4.2: *Adsorption Column Parameters*

Column Parameter	Value	Units
Cycle time	13	hours
Cycles per batch	301	–
Active adsorbent fraction at replacement	0.3	–
Mass of adsorbent	254	t/batch £
Column voidage	37	%
Number of columns	6	–
Diameter	2.68	m
Height	13.4	m

4.3 Dehumidifier

The chosen dehumidifier was the DSS3500 from Munters because it was able to process the highest flowrate of air (143,000 m^3/hr). 11 units were required to process the inlet air flowrate and an average RH of 73% (equal to the average RH of the UK) was assumed to predict the electrical costs. Note that the electrical costs will be larger for a process in a more humid climate. The results can be found in the table 4.3.

4.4 Extraction Fans

The extraction fans had to provide air at a flowrate of 1.4×10^6 m^3/hr. Using the values provided in section 3.7 of the methodology and the predicted capture of 200,000 t_{CO_2} over its 20 year lifetime, the costs of the extraction fans were calculated and are presented in table 4.4.

Table 4.3: *Cost of Dehumidifiers*

Description	Value	Units
Average Power Usage (UK)	8.96	10^6kWh/year
Average Electrical Cost	6.69	10^5 £/year
Capital Cost	2.61	10^5 £
Maintenance Cost	1.3	10^4£/year

Table 4.4: *Costs of Extraction Fans*

Cost Description	Value	Units
Capital Cost	173	10^4 £
Operational Cost	3.94	10^4 £/year
Electrical Cost	6.22	10^4 £/year

4.5 Heat Exchanger and Valves

For a heat transfer area of 24.8m^2, 175kmol/hr of cooling water and 50kmol/hr of steam was required. In this study, it is assumed that free waste heat is available. Hence, its corresponding cost is not included in the total cost calculation.

As explained in the methodology, more than 10 inlet pipes per column with a diameter of at least 0.5m would be required to achieve a reasonable Reynolds number. This is due to the CO_2 concentration in air being as low as 400ppm, resulting in extremely large flowrates of air passing through the system. Alternative solutions to avoid this will be further discussed in the section 6.1.

4.6 Economic Analysis

Table 4.5: *Costs Breakdown*

Annualised Capital Cost	
Unit	Cost (mil. £/a)
Column	0.038
Extraction Fans	0.163
Dehumidifier	0.025
Valves	0.505
Cooling Water Pump	0.007
Vacuum Pump	0.008
Heat Exchanger	0.132
Total	**0.87**
Operational Cost	
Maintenance	1.08
Utility	0.76
Raw Material Cost	
Adsorbent	25.1
Cooling Water	0.002
Total	**26.94**

In this study, an annualised capital cost of 0.87 mil. £/year was obtained, with the breakdown shown in table 4.5. The capex and the LCOD were found to be 930 £·year/t_{CO_2} and 2782 £/t_{CO_2} respectively. They will be discussed further in section 4.6 (Leeson et al. 2017b).

Over the past decade, researchers have tried to predict the costs of the first LT DAC plant. It was important to compare the calculated costs of this process model with these values in order to find discrepancies and therefore, identify areas of the process which need improvement before the first LT DAC plant can be built.

Upon comparison of the capex, it is expected that as the plant capacity increases, the capex decreases. Therefore, as can be seen from table 4.6, the capex of this process agrees extremely well with values reported by literature. However, the LCOD does not agree with the predicted values. The calculated LCOD of this process model was significantly higher than the LCOD's reported in other sources. Upon investigation of the yearly costs, it was seen in figure 4.3 that the adsorbent contributed 90% of the total cost. One possible cause of this is the scale up of raw material costs, making the cost of adsorbents less accurate as bulk purchases may have a significantly lower cost. Therefore, adsorbent cost was the primary focus of the sensitivity analysis in section 4.8.

All literature quoted in tables 4.6 and 4.7 was taken from Mahdi Fasihi (2019).

Table 4.6: *Capex*

Source	Year	Plant Capacity (t_{CO_2}/year)	Capex (£·year/t_{CO_2})
Rostenburg	2015	3,600	1037
This study	**2019**	**10,000**	**930**
Antency	2015	360,000	621

Table 4.7: *LCOD*

Source	Year	Plant Capacity (t_{CO_2}/year)	LCOD (£/t_{CO_2})
This study	**2019**	**10,000**	**2782**
Simon et al.	2011	N/A	765
Rostenburg	2015	3,600	171
Antency	2015	360,000	114
Ping et al.	2015	N/A	<96

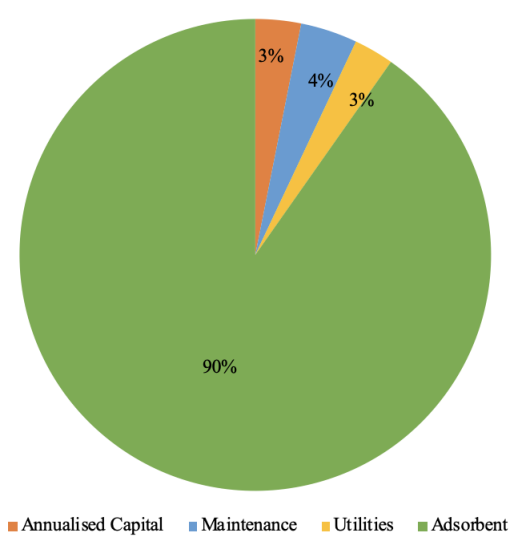


Figure 4.3: *Breakdown of Costs*

4.7 Sensitivity Analysis

4.7.1 Cost of Adsorbent

Using the cost of the raw materials, a PEI adsorbent on a HP2MGL base with a CTAB additive was predicted to have a cost of 48.3 £/kg. However, with worldwide deployment of LT DAC technology resulting in the large-scale production of this adsorbent, its cost can realistically be expected to decrease. The cost of activated carbon is roughly 5 £/kg hence this value was used as the best-case scenario for the cost of adsorbent in the sensitivity. It resulted in an 81% reduction in the LCOD of the process.

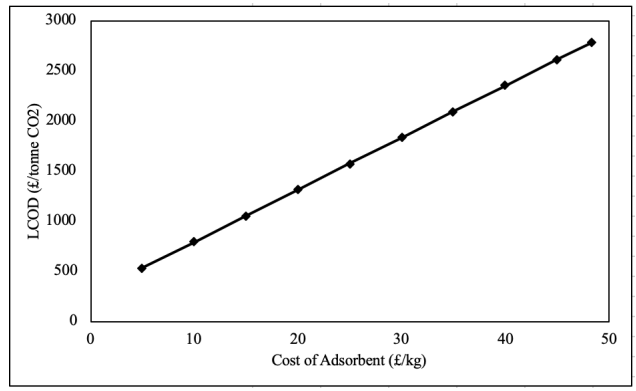


Figure 4.4: *Effect of adsorbent cost on LCOD*

4.7.2 Saturation Capacity

The saturation capacity of the adsorbent can also be expected to improve with further research on the optimal operating conditions of the column and composition of the adsorbent. With higher saturation capacities, the mass of adsorbent required will decrease, thus reducing the cost of the adsorbent, column capital and maintenance costs.

16.5 mol_{CO_2}/kg(adsorbent) was used as the upper bound for the saturation capacity as other gas adsorbents were reported to have performed this well by Hanieh Bamdad (2017). It can be seen in figure 4.5 that increasing the saturation capacity from 2.63 mol/kg to 16.5 mol/kg resulted in a 76% reduction in the LCOD and reduced the column volume by almost 65m^3.

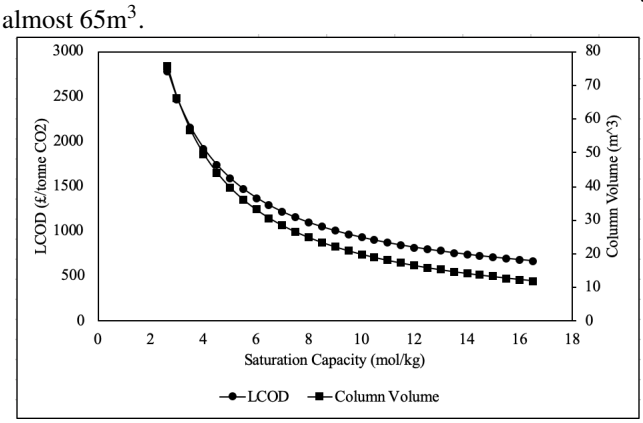


Figure 4.5: *Effect of Saturation Capacity on LCOD and Column Volume*

4.7.3 Degradation Rate

The degradation rate of the adsorbent was estimated as 0.4%/cycle by Wang et al. (2015). However, this value was only measured over 5 cycles. Given that the adsorbent used in this process was in the column for roughly 300 cycles, there was a high uncertainty in the degradation rate. It may increase due to the presence of impurities in the air or decrease upon further investigation of the adsorbent. The results of the sensitivity analysis showed that every 0.1% decrease in the degradation rate per cycle resulted in a 620 £/t_{CO_2} decrease in the LCOD.

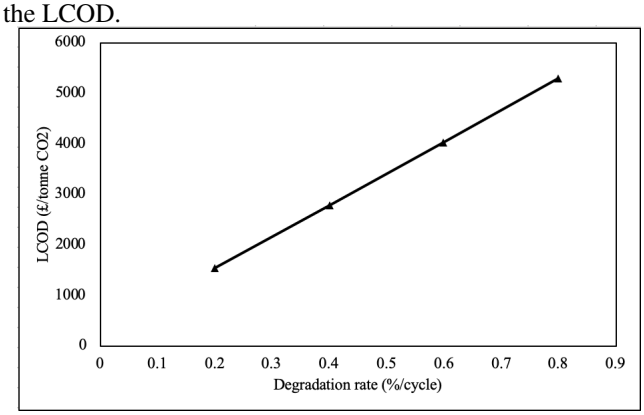


Figure 4.6: *Effect of Degradation Rate on LCOD*

4.7.4 Overview of Sensitivities

If the best case scenario of each sensitivity is taken (adsorbent cost = 5 £/kg, saturation capacity = 16.5 mol/kg and degradation rate = 0.2 %/cycle), then the LCOD of the process reduces to just 293 £/t_{CO_2}. This is comparable to the values reported in literature and therefore would suggest that the disparity was largely due to cost and properties of the adsorbent used in the study.

5 Downstream Utilisation

Table 5.1: *Transport*

Transport type	Capacity	Distance (Mt_{CO_2}/a)	Cost (£/t_{CO_2})
Train	1.46	598	6.1
Onshore Pipeline	0.73	100	5.7
Offshore pipeline	2.5	180	7.8
Shipping	2	750	9.3

Two possible downstream utilisations will be discussed in this section, methanol production and carbon storage. For both cases, the cost of transporting CO_2 must be taken into account.

Depending on the type of transport, capacity of CO_2 being transported, as well as travel distance, an estimation of the cost of transport can be obtained, as summarised by Mahdi Fasihi (2019) and shown in table 5.1. The cost of transport would add a maximum of 5% to the predicted LCOD from the best-case scenario found in the sensitivity analysis. This is not an insignificant proportion of the LCOD and therefore this DAC process should be located close to the downstream utilisation of the extracted CO_2 to minimise cost.

5.1 Methanol Production

CO_2 can be used as feed for methanol production. However, different conclusions have been made about the demand for CO_2 as feed. Konstantinos Atsonios (2016) explained that compared to conventional ways of producing methanol, using CO_2 as feed gives a much higher cost, deterring the uptake of this technology. Mar Pérez-Fortes (2016) claims that there is a good market demand for CO_2 as feed in Europe, around 41Mt_{CO_2}/year. Research papers on this type of methanol production, for example, that studied by Prapatsorn Borisut (2019) mostly assume pure CO_2 feed, hence further purification of captured CO_2 must be considered, which implies an increase in cost. Therefore, further research must be carried out to ensure that demand is present before performing purification.

5.2 Carbon Storage

If the captured carbon dioxide cannot be used to produce methanol, then an alternative solution is to store it underground. Sites for injection include coalbeds, deep saline aquifers and depleted oil and gas reservoirs with their estimated storage capacity of CO_2 falling between 8,000 and 55,000 Gt_{CO_2} (Kearns et al. 2017). Care must be taken during the early, injection stage to ensure it does not leak and pollute local water sources. However, once stored it is unlikely to leak.

The costs associated with carbon storage are shown in the table 5.2 (Howard Herzog 2005). It can be seen that the cost of storage is roughly 2% of the total LCOD found in the sensitivity.

Table 5.2: *Cost of Storage*

Cost £/t_{CO_2}	Lower Bound	Upper Bound
Geological Storage	0.5	8.1
Geological Monitoring	0.1	0.3

6 Conclusion

An LT DAC plant was designed to capture 10,000 t_{CO_2}/year using the PEI adsorbent with a HP2MGL support and a CTAB diffusion additive. An LCOD of approximately £293/t_{CO_2} can be achieved with improvements to the adsorbent's cost, degradation rate and adsorption capacity. A Capex of 930 $£{\cdot}year/t_{CO_2}$ was also obtained. Although other large-scale CO_2-removal techniques such as BECCS, Reforestation, and Enhanced Ocean Productivity exist, they all have major environmental impacts such as affecting land or marine life. Since the structure of a DAC plant is similar to that of a wind turbine, the land in between columns can be used for other purposes such as agricultural activities, hence eliminating any

major environmental impacts. For this reason, DAC is proposed as the most viable option for large-scale removal of CO_2.

However, in order to encourage large up-take of DAC to meet the Paris Agreement, further efforts would need to be made. Some examples include more government subsidies, increase in carbon pricing as well as increase in demand for downstream utilisation. With these, countries can increase their prospect of paving the way to a sustainable future.

6.1 Alternative Solution

Cryogenic separation is a possible alternative as it removes up to 96vol% of the inlet air before entering the process, hence reducing the subsequent equipment sizes and reducing costs. Since CO_2 has the second highest boiling point in air, with a large disparity from the component with the next highest boiling point, it can be removed by using molecular sieves, avoiding high cost of cryogenic distillation. In the study conducted by Amarkhail (2009), the air stream was compressed to approximately 6 bar and cooled to 25°C using inter-stage cooling, before the condensed water and CO_2 was removed. This increases the concentration of CO_2 in the inlet stream greatly, potentially resulting in a better adsorption capacity. However, the complicated behaviour of adsorption capacity towards CO_2 in steam rather than air is yet to be understood well so this needs to be taken into account when considering this method. Finally, although the cost of subsequent equipment may be reduced, the cost of compression and inter stage cooling is still comparable to this study. A simple simulation in ASPEN Plus using short-cut models for heat exchangers was carried out which gave a capital cost of approximately £35milion and utility cost of £8million. This can be attributed to the high flowrate of air that is being compressed as well as the high temperature air that needs to be cooled down to ambient temperatures.

6.2 Further Efforts

In this section, further efforts required to continue what was set out to do in this paper, as well as the efforts required to encourage wide uptake of DAC.

The following improvements can be made to this study:

- **Modelling using Aspen Adsorption** – following further research on the adsorbent, if the physical and isotherm parameters are found then Aspen Adsorption can be used to model the adsorption column

- **Compression of the inlet air** – another utilisation of the isotherm parameters of the adsorbent in this study would be the ability to compress the inlet air. Although compressors would have to be added to the process increasing the overall capital and electrical costs, this would allow the adsorption column to be designed at higher pressures. A higher saturation capacity of the adsorbent could be achieved thus reducing the column volume and adsorbent costs

- **Accuracy** – in order to improve the accuracy of this model, information regarding the degradation rate of the adsorbent measured over many cycles and the cost of the adsorbent in bulk would be required

To encourage wide uptake of DAC of CO_2 in order to bring us closer to meeting the Paris Agreement, government subsidies must be in place. Carbon pricing must also show a steady increase to act as an incentive. Demand for possible downstream utilisation of captured CO_2 must also be studied, as already mentioned in section 5.

7 Acknowledgements

The authors would like to thank Sara Giarola from the Sustainable Gas Institute (SGI), for providing guidance and advice throughout this project. We would also like to thank Dr. Colin Hale, for giving us advice whenever we faced difficulties during the design of our equipment.

Nomenclature	Description	Units
d_r	Degradation rate	-
f	Fraction of Active Adsorbent at Replacement	-
$m_{CO_2,batch}$	Carbon Capture per batch	t_{CO_2}/batch
$\dot{m}_{CO_2}$	Carbon Capture per year	t_{CO_2}/year
n	Cycle Number	-
Cap_n	Saturation Capacity at Cycle n	$mol_{CO_2}/kg_{adsorbent}$
$m_{CO_2,ads}$	Carbon Capture per Mass of Adsorbent	$mol_{CO_2}/kg_{adsorbent}$
N	Total Number of Cycles	-
$m_{ads,batch}$	Mass of Adsorbent per batch	$kg_{adsorbent}$/batch
NC	Number of Columns	-
V	Individual Column Volume	m^3
ρ	Adsorbent Density	kg/m^3
ε	Column Voidage	-
D	Diameter	m
H	Height	m
P	Pressure	kPa
F_{mat}	Material Cost Factor	-
$M_{air,col}$	Mass of Residual Air in the Column	kg_{air}
C_{CO_2}	Carbon Capture per Cycle	kg_{CO_2}/cycle
crf	Annuity Factor	/year
l	Plant Lifetime	years
ACC	Annualised Capital Cost	£/year
$Capex$	Capital Expenditure	£/t_{CO_2} $\cdot year$
OC	Operational Cost	£/year

References

Amarkhail, Sher Shah (2009). "Air Separation". In: URL: http://kchbi.chtf.stuba.sk/ODAOld/doc/Diploma%5C%20work%5C%20of%5C%20Shershah%5C%20Amarkhail.pdf.

B. Wetenhall J.M. Race, M.J. Downie (2014). "The Effect of CO_2 Purity on the Development of Pipeline Networks or Carbon Capture and Storage Schemes". In: *International Journal of Greenhouse Gas Control* 30, pp. 197–211. DOI: http://dx.doi.org/10.1016/j.ijggc.2014.09.016.

David W. Keith Minh Ha-Duong, Joshuah K. Stolaroff (2005). "Climate Strategy with CO_2 Capture from the Air". In: *Climate Change* 74, pp. 17–45. DOI: https://doi.org/10.1007/s10584-005-9026-x.

E.Brunazzi G. Nardini, A. Paglianti (2002). "An Economical Criterion for Packed Adsorption Column Design". In: *Chem. Biochem. Eng. Q.* 4, pp. 199–206. URL: http://pierre.fkit.hr/hdki/cabeq/pdf/16_4_2002/Cabeq_2002-04_1.pdf.

Flomatic Valves Company (n.d.). *2019 Price List*. URL: https://www.lesker.com/newweb/vacuum_pumps/screwpump_busch_cobra.cfm.

Hanieh Bamdad Kelly Hamboldt, Stephanie MacQuarrie (2017). "A Review on Common Adsorbents for Acid Gases Removal: Focus on Biochar". In: *Renewables and Sustainable Energy Reviews*. DOI: 10.1016/j.rser.2017.05.261.

Heidel, Kenton et al. (2011). "Process Design and Costing of an air-contactor for air capture". In: *Energy Procedia* 4, pp. 2861–2868. DOI: https://doi.org/10.1016/j.egypro.2011.02.192.

Howard Herzog, Koen Smekens (2005). "Special Report on Carbon Dioxide Capture and Storage". In: *IPCC*, pp. 341–360. URL: https://www.ipcc.ch/site/assets/uploads/2018/03/srccs_chapter8-1.pdf.

IEA (2012). "Energy Technology Perspectives :Pathways to a Clean Energy System". In: DOI: https://doi.org/10.1787/energy_tech-2012-en..

— (2017). "CO_2 Emissions for Fuel Combustion". In: URL: https://www.iea.org/publications/freepublications/publication/%20CO2EmissionsfromFuelCombustionHighlights2017.pdf..

— (2018). "CO2 emissions from fuel combustion Journal: Secure Sustainable Together". In: URL: http://www.indiaenvironmentportal.org.in/files/file/CO2_Emissions_from_Fuel_Combustion_2018_Highlights.pdf.

Kearns, J. et al. (2017). "Developing a Consistent Database for Regional Geologic CO2 Storage Capacity Worldwide". In: *Energy Procedia* 114, pp. 4697–4709. DOI: 10.3389/fenrg.2019.00081.

Konstantinos Atsonios Kyriakos D. Panopoulos, Emmanuel Kakaras (2016). "Investigation of technical and economic aspects for methanol production through CO_2 hydrogenation". In: *International Journal of Hydrogen Energy* 4,

pp. 2202–2214. DOI: http://dx.doi.org/10.1016/j.ijhydene.2015.12.074.

Kurt K. Lesker Company (n.d.). *Busch Cobra Series Screw Vacuum Pump*. URL: https://www.lesker.com/newweb/vacuum_pumps/screwpump_busch_cobra.cfm.

Kurt Zenz House, et al. (2011). "Economic and energetic analysis of capturing CO_2 from ambient air". In: *PNAS* 108.51, pp. 20428–20433. DOI: www.pnas.org/cgi/doi/10.1073/pnas.1012253108.

Lackner, K.S. (2009). "Capture of carbon dioxide from ambient air". In: *The European Physical Journal Special Topics* 176, pp. 93–106. DOI: https://doi.org/10.1140/epjst/e2009-01150-3.

Leeson, D. et al. (2017a). "A Techno-economic analysis and systematic review of carbon capture and storage (CCS) applied to the iron and steel, cement, oil refining and pulp and paper industries, as well as other high purity sources". In: *International Journal of Greenhouse Control* 61, pp. 71–84. DOI: https://doi.org/10.1016/j.ijggc.2017.03.020.

— (2017b). "A Techno-economic analysis and systematic review of carbon capture and storage (CCS) applied to the iron and steel, cement, oil refining and pulp and paper industries, as well as other high purity sources". In: *International Journal of Greenhouse Control* 61, pp. 71–84. DOI: https://doi.org/10.1016/j.ijggc.2017.03.020.

Mahdi Fasihi Olga Efimova, Christian Breyer (2019). "Techno-economic assessment of CO_2 direct air capture plants". In: *Journal of Cleaner Production* 224, pp. 957–980. DOI: https://doi.org/10.1016/j.jclepro.2019.03.086.

Mar Pérez-Fortes, Evangelos Tzimas (2016). "Techno-economic and Environmental Evaluation of CO_2 Utilisation for Fuel Production: Synthesis of Methanol and Formic Acid". In: *JRC Science for Policy Report*. DOI: 10.2790/981669.

Maryam Mahmoudkhani, David W.Keith (2009). "Low-energy sodium hydroxide recovery for CO_2 capture from atmospheric air—Thermodynamic analysis". In: *International Journal of Greenhouse Gas Control* 3, pp. 376–384. DOI: https://doi.org/10.1016/j.ijggc.2009.02.003.

Micah Broehm Jessica Strefler, Nico Bauer (2015). "Techno-Economic Review of Direct Air Capture Systems for Large Scale Mitigation of Atmospheric CO_2". In: *SSRN*. DOI: http://dx.doi.org/10.2139/ssrn.2665702.

Prapatsorn Borisut, Aroonsri Nuchitprasittichai (2019). "Methanol Production via CO_2 Hydrogenation: Sensitivity Analysis and Simulation—Based Optimization". In: *Frontiers in Energy Research* 7. DOI: 10.3389/fenrg.2019.00081.

Realmonte, Giulia et al. (2019). "An inter-model assessment of the role of direct air capture in deep mitigation pathways". In: *Nature Communications* 10. DOI: https://doi.org/10.1038/s41467-019-10842-5.

Simon, A.J. et al. (2011). "Systems analysis and cost estimates for large scale capture of carbon dioxide from air". In: *Energy Procedia* 4, pp. 2893–2900. DOI: https://doi.org/10.1016/j.egypro.2011.02.196.

Sinha, Anshuman et al. (2017). "Systems Design and Economic Analysis of Direct Air Capture of CO_2 through Temperature Vacuum Swing Adsorption Using MIL-101(Cr)-PEI-800 and mmen-Mg2(dobpdc) MOF Adsorbents". In: *Industrial & Engineering Chemistry Research* 3.56, pp. 750–764. DOI: DOI:10.1021/acs.iecr.6b03887.

W.Keith, David et al. (2018). "A Process for Capturing CO2 from the Atmosphere". In: *Joule* 2, pp. 1573–1594. DOI: https://doi.org/10.1016/j.joule.2018.05.006.

Wang, Jitong et al. (2015). "Application of polyethylenimine-impregnated solid adsorbents for direct capture of low-concentration CO_2". In: *AIChE Journal* 61, pp. 972–980. DOI: https://doi.org/10.1002/aic.14679.

Williamson, Phil (2016). "Scrutinise CO_2 Removal Methods". In: *Nature* 530, pp. 153–155. URL: https://www.nature.com/articles/530153a.pdf.

ATR-FTIR Imaging Technique for Study of Factors Influencing CBZ-NIC Cocrystal Tablet Dissolution

Yian-Shin Lai and Cheow Lin Tan
Department of Chemical Engineering, Imperial College London, U.K.

Abstract

Cocrystallisation was used to improve the bioavailability of poorly soluble carbamazepine (CBZ), the active pharmaceutical ingredient (API) in this study by adding nicotinamide (NIC) as a coformer. Crystallisation of CBZ into dihydrate occured, which significantly decreased the dissolution rate of the API. The effects of using different excipients in each tablet formulation; hydroxypropyl cellulose (HPC), hydroxypropyl methylcellulose (HPMC), polyethylene glycol 4000 (PEG) and saccharin (SAC) were explored using attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopic imaging. The dynamic dissolution of tablets was studied and strong effects from changing the excipients were observed. Cocrystal-HPC tablet exhibited the fastest dissolution followed by cocrystal-HPMC, cocrystal-PEG, and cocrystal-SAC tablets. Dissolution rate showed a strong correlation to the amount of CBZ dihydrate formation. Results indicated HPC as the most promising polymer to achieve a controlled release of CBZ. pH of dissolution medium did not show a strong influence on tablet dissolution rates, and this was further supported by acid-base chemistry knowledge.

Keywords: *ATR-FTIR, imaging, cocrystal, CBZ, NIC, dissolution*

1. Introduction

1.1. Formulation

Oral ingestion of tablets is the preferred route of administration because of convenience and economic advantages. Tablets are typically compacted powder mixtures of active pharmaceutical ingredients (API), which deliver some therapeutic effect and pharmacologically inert excipients that bulk up formulations and facilitate drug absorption. The relevance of analysing tablet dissolutions in different pH mediums grows in importance as oral formulations pass through the gastrointestinal tract, which has different pH in different segments before being absorbed into the bloodstream (Homayun, et al., 2019). The gastrointestinal tract pH range is from 1.2 to 7.4 (Sugano & Terada, 2015). The Biopharmaceutics Classification System (BCS) pioneered by Lennernäs, et al. (1995) and currently employed by the US Food and Drug Administration (FDA) is commonly used to categorise drugs based on aqueous solubility and permeability. BCS class II and IV drugs have low solubilities and about 40% of marketed drugs and 90% of undeveloped drugs fall under these categories, causing low drug efficacies (Loftsson & Brewster, 2010). Thus, there is ongoing active research to improve the bioavailability of these drugs employing prodrugs and drug conjugates, medicinal chemistry and formulation design approaches (Gomez-Orellana, 2005).

A method gaining traction to increase the solubility of BCS class II drugs is cocrystallisation (Emami, 2018). Cocrystals are crystalline materials consisting of API and cocrystal formers (coformers) following a stoichiometric ratio in the same lattice. The molecules are typically connected by non-covalent interactions such as hydrogen bonds, halogen bonds, π-π interactions and van der Waals' forces (U.S. Department of Health and Human Services FDA, 2018; Cao, et al., 2018). Cocrystallisation is shown to modify physicochemical properties of APIs such as dissolution rate, aqueous solubility and melting point without compromising structural integrity (Sathisaran & Dalvi, 2018; Schultheiss & Newman, 2009). Coformer selection is a vital process in achieving these desired modifications and normally uses either a trial-and-error method from screening a library of 'generally recognised as safe' (GRAS) status or a supramolecular synthon method (Fukte, et al., 2014).

A great model drug in cocrystal research is carbamazepine (CBZ) which is largely used as an antiepileptic drug to treat convulsions and sometimes as a mood stabilizer for bipolar disorder or to treat nerve pains. CBZ falls under BCS class II but it contains an amide functional group that allows interaction with

a wide range of coformers (Fleischman, et al., 2003). CBZ cocrystals demonstrated better dissolution rates compared to their amorphous forms in Drozd, et al.'s paper (2017) and Sathisaran and Dalvi's research (2019). Nicotinamide (NIC) is widely used as an effective coformer in enhancing intrinsic dissolution rates due to its high solubility (Yuliandra, et al., 2018; Wang, et al., 2013; Shewale, et al., 2015) and is GRAS by the FDA. Hence, this study will be employing cocrystals with CBZ as the API and NIC as the coformer. CBZ exists as polymorphs and pseudopolymorphs (Kobayashi, et al., 2000); a particular form of interest is CBZ dihydrate which is readily formed from dissolving CBZ in aqueous solution through 'solution mediated phase transformation' (Greco & Bogner, 2012). This can lead to a significant decrease in dissolution rate of the API. Many researchers have proposed the addition of excipients to tablet formulations as a method to inhibit the formation of CBZ dihydrate and there is growing interest in understanding the mechanism of drug dissolution of cocrystal-excipient matrices (Huang & Rodríguez-Hornedo, 2010; Qiu & MingzhongLi, 2015; Remenar, et al., 2007). The excipients selected for this experiment are polyethylene glycol 4000 (PEG), saccharin (SAC), hydroxypropyl cellulose (HPC) and hydroxypropyl methylcellulose (HPMC) and their effects on the API dissolution rate were explored.

1.2. Dissolution

United States Pharmacopeia (USP) dissolution tests has been widely adopted in the pharmaceutical industry to obtain drug release profiles and assist in quality control (Cohen, et al., 1990). It usually involves dissolving a drug in a vessel containing solution and quantifying the amount of drug released with a UV/Vis spectrometer. General chapter <711> in the USP (The United States Pharmacopeial Convention, 2011) describes four main apparatus in which Apparatus 1 utilising a basket stirring element was chosen in this experiment due to its suitability for solid oral formulations. However, this USP testing method is unable to fully characterise drug dissolution as it provides only quantitative data and shows large variability in results (D'Arcy, et al., 2006). Thus, there is a need to supplement this test with other methods to obtain more qualitative data.

Various spectroscopic imaging techniques such as magnetic resonance imaging (MRI), Fourier transform infrared spectroscopy (FTIR) and Raman spectroscopy and mapping become increasingly useful in providing complementary information from in-situ monitoring and observation of drug dissolutions (Punčochová, et al., 2017; Zhang, et al., 2011; Qiao, et al., 2013). Among these techniques, FTIR spectroscopic imaging was utilised in this study. There are three methods of FTIR commonly used: transmission FTIR, attenuated total reflection FTIR (ATR-FTIR) and diffuse reflection infrared Fourier transform (DRIFT). Transmission and DRIFT are not suitable for the study of samples in an aqueous medium because water strongly absorbs infrared and will greatly attenuate the incident signal reaching the detector causing spectra saturation. On the other hand, ATR-FTIR spectroscopy is applicable in aqueous mediums and is useful for observing a small part of tablets during in-situ dissolution experiments, making it particularly relevant to this experiment and hence, it was used throughout this study.

The setup of ATR-FTIR mode involves a highly refractive prism element or crystal to be in contact with a sample surface and the basic principle utilised is total internal reflection. On such basis, the ATR crystal must have a higher refractive index than the sample studied, and examples of ATR crystals are diamond and zinc selenide. The diamond ATR crystal was used in this experiment due to its superior mechanical strength and chemically inert properties. The angle of the incident radiation must exceed the critical angle to achieve total internal reflection and the relationship is written as follows:

$$\theta_c = sin^{-1}\left(\frac{n_2}{n_1}\right)$$

where θ_c is the critical angle, n_1 and n_2 is the refractive indices of the ATR crystal and the sample respectively.

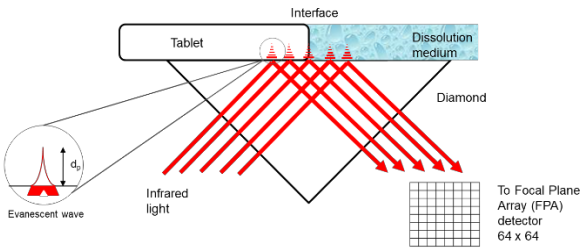


Figure 1 Schematic of ATR-FTIR spectroscopic imaging

Figure 1 shows the working principle of ATR-FTIR spectroscopic imaging. When the incident infrared radiation shines on the diamond and total internal reflection occurs, some of the radiation probes into the sample by a few micrometers, forming an evanescent wave with an exponentially decreasing magnitude. The depth of penetration is the depth of evanescent wave that extends into the sample until its

amplitude falls to e^{-1} of its surface value and can be calculated as follows:

$$d_p = \frac{\lambda}{2\pi n_1 \left(sin^2\theta - \left(\frac{n_2}{n_1}\right)^2 \right)^{1/2}}$$

where d_p is the depth of penetration, λ is the wavelength of radiation, θ is the angle of incidence, n_1 and n_2 are the refractive indices of the ATR crystal and the sample respectively (Kazarian & Ewing, 2013). The depth of penetration depends on wavelength, angle of incidence and refractive index of ATR crystal and medium. It is very small hence only allowing only a very thin layer of the tablet to be scanned. Due to this, good contact between the sample and diamond surface is importance to prevent any trapped air or medium and obtain clear spectra. After hitting the diamond, the beam is then directed by mirrors to a Focal Plane Array (FPA) detector and the absorbance spectra obtained can be integrated and combined to form a 64 by 64 pixel 'chemical image'.

The goal of this experiment was to understand the dissolution of cocrystal tablets made with different excipients under different pH conditions using ATR-FTIR spectroscopic imaging and deduce factors that can govern a controlled release of CBZ.

2. Methodology

2.1. Tablet Formulation

Carbamazepine (CBZ), Nicotinamide (NIC), and Saccharin (SAC) were purchased from Sigma-Aldrich. Polyethylene glycol 4000 (PEG) was purchased from Fischer Chemicals. Hydroxypropyl Methylcellulose K100 Premium LV CR EP (HPMC) was supplied from Colorcon. The source of hydroxypropylcellulose (HPC) was not identified but the substance was tested using FTIR spectroscopy and the resultant spectra obtained was compared with HPC spectra in literature . Pure cocrystal tablets were made from powder CBZ and NIC using a 1:1 molar ratio. The CBZ and NIC weighed were placed in a ceramic mortar and liquid-assisted grinding was carried out using a ceramic pestle (Karimi-Jafari, et al., 2018). Both powders were grinded together while dropping ethanol every few minutes in the mixture for a duration of 30 minutes. The cocrystal powder formed was then placed in a sample vial for storage at room temperature. The choices of excipients encompass SAC, PEG, HPMC and HPC. Cocrystal powder and excipient were combined in 80:20 by weight and grinded together for 5 minutes to reduce the size of the powder particles and promote further mixing. Each excipient mixture was then stored in a separate sample vial at room temperature.

Figure 2 Metal dissolution cell accessory

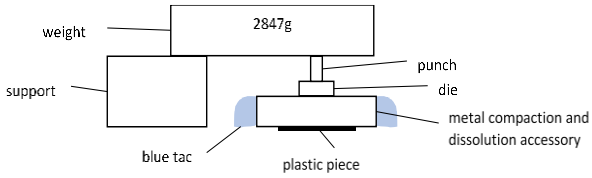


Figure 3 Schematic of tablet formulation setup

An unconventional tablet compaction method was used due to equipment unavailability in the laboratory. A metal dissolution cell accessory as shown in Figure 2 was used which consisted of an external cell, a die, and a punch. The cell was first secured using blue tack with the cell extrusion lying on top of a clean flat firm plastic piece cut out from weighing boats. The die was then screwed in until the bottom surface of the die lay flat on the plastic. Next, 10 mg of tablet powder mixture was weighed out and poured into the die extrusion using a paper funnel. To standardise the compaction force, a metal weight of 2847 g was placed on the punch for 2 minutes and assisted by a support to allow for the weight to be balanced laying flat on the punch. After this process, the die was slowly unscrewed, and the punch was pushed to release the compacted tablet. The whole setup was then dissembled, cleaned and the whole procedure was repeated for each tablet made.

2.2. Acidic Dissolution Medium Preparation

Experiments were carried out with distilled (DI) water and solutions of pH4 and pH5. Solution was adjusted to respective pH values by dropwise addition of HCl (0.1M) and sodium hydroxide (0.1M).

2.3. USP Dissolution

USP type I method for dissolution of tablets was carried out with the Vankel VK7000 Dissolution system. Tablets were put in standard-compliant stainless steel baskets, and solutions of 300 ml were filled in each vessel. The temperature of the water bath was set to 25˚C to allows for comparability with FTIR imaging results and experiments were carried out after

temperature equilibrium. The baskets were spun at a rotational speed of 250 rpm for a duration of 2 hours 35 minutes (Note: the process does not follow the standard 100 rpm and did not reach full completion as per FDA requirements due to time constraints; results were compared relative to each other but cannot be used to determine absolute dissolution rate). Three samples were drawn out manually from each vessel every 15 minutes and analysed under a PerkinElmer Lambda35 UV/Vis Spectrometer. Quartz cuvettes were used as the range of absorbance wavelength lies in the UV region (10 nm-400 nm).

Radiation of wavelength 300 nm was selected to measure the absorbance of CBZ. It was shown by Qiao, et al. (2013) that UV absorbance of NIC contributes less than 5% at 300 nm for cocrystal of CBZ:NIC at a 1:1 ratio. Spectra of the cocrystal and each excipient were also collected to examine the contribution of each excipient at 300 nm. Excipient contributions were found to be negligible due to their low concentrations in the formulations.

At the end of each experiment, all undissolved tablets were fully dissolved using an ultrasonic water bath, and the final absorbance was obtained on the spectrometer. A graph of percentage of CBZ dissolved against dissolution time was then plotted for analysis.

2.4. Tablet Dissolution Set Up

Tensor 27 FTIR spectrometer (Brucker, Germany) was coupled with an IMAC macro imaging chamber and a 64×64 FPA detector as shown in Figure 4. These are used for macro FTIR imaging to obtain 4096 spectra in each scan. Images of size 0.64 mm × 0.53 mm with a spatial resolution of 10–15 µm (Ewing, et al., 2015) were produced. 64 scans were taken for each measurement which takes about 4.5 minutes and the spectral resolution was obtained at 4 cm^{-1}.

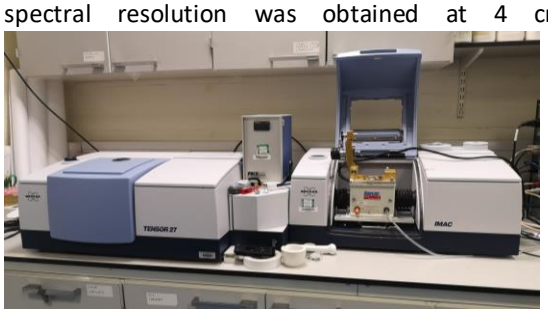


Figure 4 Tensor 27 FTIR spectrometer and IMAC macro imaging chamber

Measurements were taken every five minutes and analysed with OPUS. The parameters selected to obtain data for the whole experiment were gain of 1, video offset of 245 and exposure time of 0.212125 ms.

A customized Specac imaging Golden gate was used with a 2×2 mm^2 diamond crystal. Each tablet was placed on the diamond crystal and adjusted for an interface to be observed before each experiment. A polyacrylic cell was used to flow through solvent pumped by the Original-Perfusor-Spritze VI pump. Flowrate of solvent was set to 60 ml/hr on the pump.

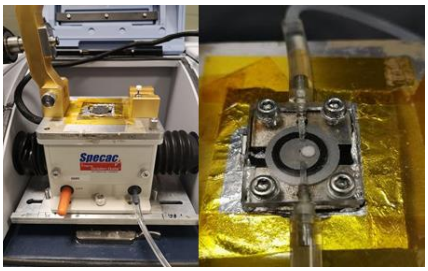


Figure 5 Golden gate (Specac) and polyacrylic flow cell setup

Ten sets of dissolution tests were carried out for four different cocrystal-excipent combinations in three pH mediums, water and acidic (pH4, pH5). Cocrystal-SAC dissolution tests in acids were not carried out due to time constraints. Each experiment was continued for a duration of two hours and repeated to verify the consistency of the results.

2.5. Preliminary Chemical Spectra Collection and Integration Method

To analyse the distribution of a specific substance in the matrix, characteristic bands were identified. This was carried out with Alpha-P (Brucker), a single reflection diamond ATR-FTIR spectrometer. 64 scans were taken for each measurement with a spectral resolution of 4 cm^{-1}. Figure 6 shows the spectra of CBZ-NIC cocrystal, and CBZ dihydrate. The spectra were compared to each other and characteristic integration peaks for cocrystal and CBZ dihydrate were determined to be from 1670 cm^{-1} to 1645 cm^{-1}, and 1574 cm^{-1} to 1550 cm^{-1} respectively. The presence of the cocrystal can be verified by the strong characteristic peaks at 1683 cm^{-1} and 1658 cm^{-1} which results from the C=O stretching of CBZ and NIC respectively (Qiao, et al., 2013). Additionally, both spectra were compared to those of excipients to ensure no overlap of the peaks occurred in the chosen range.

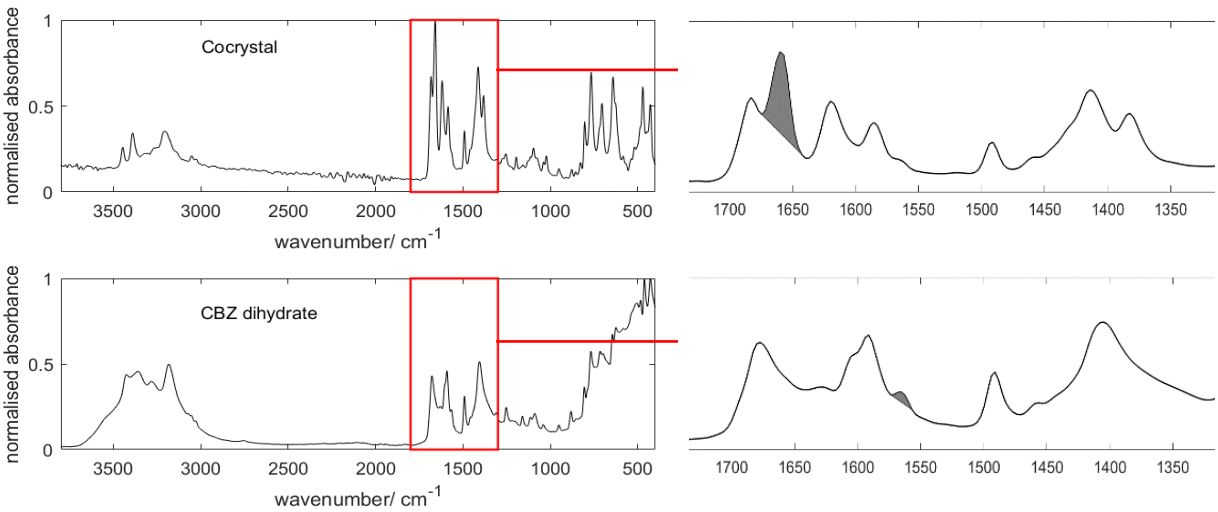


Figure 6 Single channel spectra of cocrystal and CBZ dihydrate

Tensor was then used for the dissolution experiment. The spectra obtained were integrated at the determined band range for each substance, allowing the area under the graph to be calculated to produce a heat map of the substance from low concentration to high concentration as demonstrated in Figure 7 and Figure 8. A clear interface was shown between the cocrystal tablet and solution. The substance present at different times of the experiment can be clearly observed.

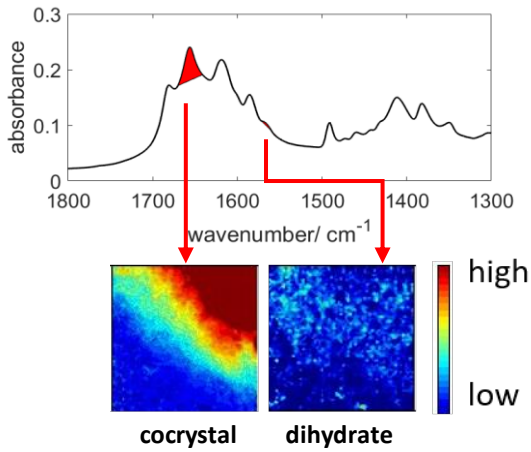


Figure 7 Cocystal-PEG spectrum and images obtained at time 0 of dissolution

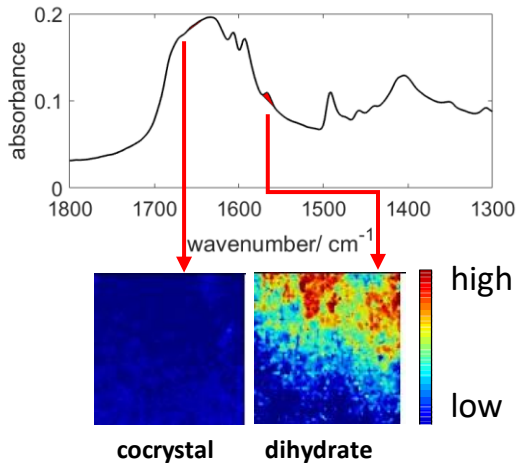


Figure 8 Cocrystal-PEG spectrum and images obtained after 1 hour of dissolution

3. Results and discussion

3.1. USP testing

The graph in Figure 9 shows the dissolution profiles of eight tablets made from four different excipients dissolved in water and pH4 acid. Dissolution in pH5 acid was not plotted because similar trends to the pH4 graphs were exhibited.

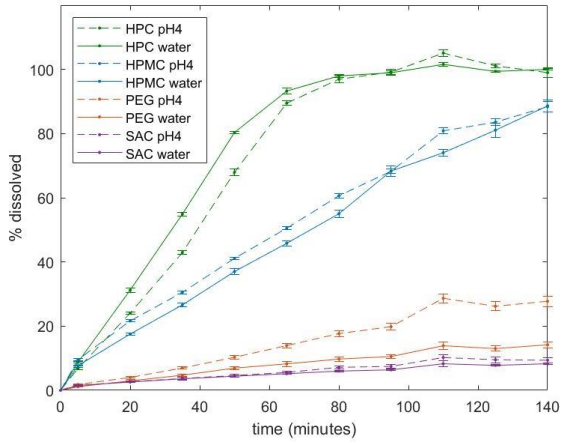


Figure 9 Percentage of CBZ dissolved against time for cocrystal tablets made with different excipients dissolved in pH4 acid and DI water

3.1.1. Effects of pH on dissolution

The pH of the dissolution environment appears to show a slight effect on the rates of dissolution but there is actually no discernible trend observed. There is poor correlation of dissolution medium pH to drug dissolution rate which can be deduced from the crossing of the acid and water graph lines of cocrystal-HPC tablets and cocrystal-HPMC tablets. While there seems to be a trend for cocrystal-PEG tablets, this result will be contradicted by a series of ATR-FTIR images in Section 3.2.1.

The reason pH of dissolution medium shows a weak correlation to tablet dissolution rate can be explained using chemistry. The pKa values are found to be -3.8 and 15.96 for CBZ (Drugbank, 2005) and 3.35 and 13.39 for NIC (U.S. National Library of Medicine, 1984). Thus, no protonation would occur in a pH4 aqueous environment. Further experiments under more acidic conditions need to be carried out to observe a change in dissolution rate. As the most basic site available on NIC is the nitrogen in the pyridine ring as shown in Figure 10b, protonation only occurs on that site and does not affect the hydrogen bonds that make up the CBZ-NIC cocrystal hence, pH is expected to show no effect on dissolution rate of the cocrystal tablets unless it is reduced further. Perhaps a change in the pH environment may affect the excipients, leading to small variations in the results. Hence, more repetitions need to be carried out to obtain a more definitive trend.

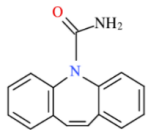


Figure 10a Carbamazepine

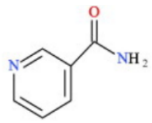


Figure 10b Nicotinamide

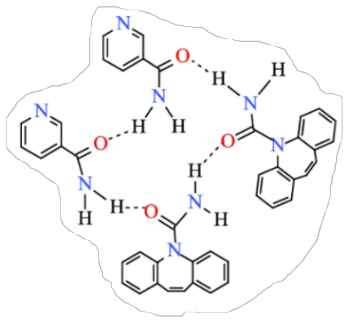


Figure 10c Cocrystal structure and Hydrogen Bonding (Qiao, et al., 2013)

3.1.2. Effects of different excipients on dissolution

Among all the tablets, only the cocrystal-HPC tablet fully dissolves under the duration of the experiment in both water and acid dissolution mediums which can be observed from the graph lines that plateau at the start of 80 minutes. This indicates that the cocrystal-HPC tablet dissolves with the highest rate. The next steepest gradient belongs to the graphs of cocrystal-HPMC tablet followed by cocrystal-PEG tablets and cocrystal-SAC tablets.

3.2. ATR-FTIR imaging

3.2.1. Effects of pH on dissolution

Comparing Figure 11 and Figure 12, more dihydrate has formed for cocrystal-PEG tablets in acidic conditions and hence, cocrystal-PEG is deduced to exhibit slower dissolution. This however contradicted the USP results, where cocrystal-PEG tablets were dissolving at a larger rate in acidic conditions, proving the insignificance of pH on dissolution rate. Both cocrystal-HPC and cocrystal-HPMC showed similar dissolution rates in water and acidic conditions.

3.2.2. Effects of different excipients on dissolution

The following correspond to ATR-FTIR images obtained from the tablet dissolution experiment grouped by excipient used and pH of dissolution medium. All the pictures in the first row correspond to the cocrystal (integrated from 1670 cm^{-1} to 1645 cm^{-1}) and all the pictures in the second row correspond to the dihydrate (integrated from 1574cm^{-1} to 1550cm^{-1}).

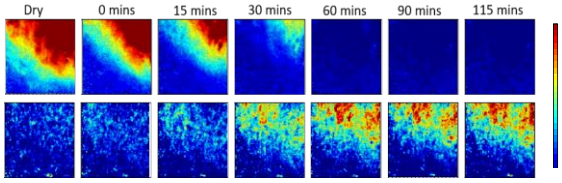


Figure 11 Cocrystal-PEG tablet dissolution in water

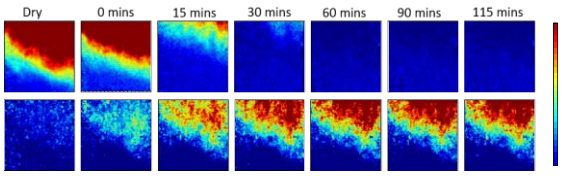


Figure 12 Cocrystal-PEG tablet dissolution in pH4 acid

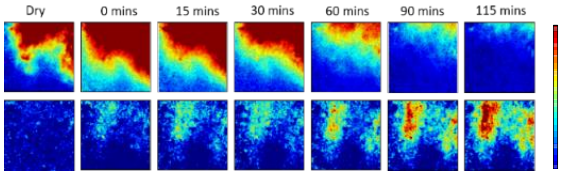


Figure 13 Cocrystal-HPMC tablet dissolution in water

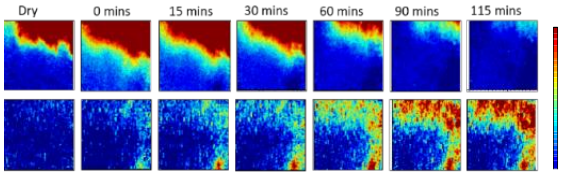


Figure 14 Cocrystal-HPMC tablet dissolution in pH4 acid

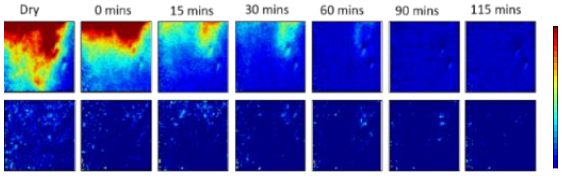


Figure 15 Cocrystal-HPC tablet dissolution in water

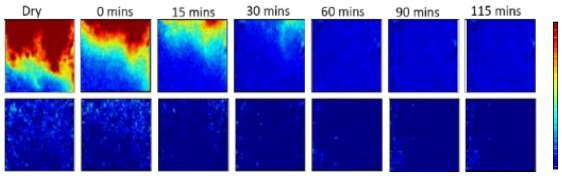


Figure 16 Cocrystal-HPC tablet dissolution in pH4 acid

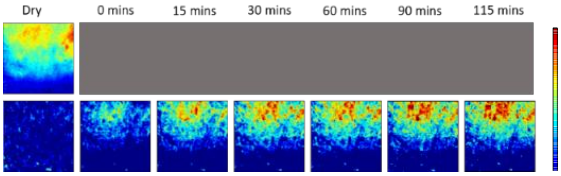


Figure 17 Cocrystal-SAC tablet dissolution in water

3.2.1.1. Polyethylene glycol 4000 (PEG) excipient

Figure 11 and Figure 12 showed a relatively fast decrease in cocrystal area for cocrystal-PEG tablets compared to both cocrystal-HPC and cocrystal-HPMC tablets, disappearing in around 30 minutes. However, an almost instantaneous dihydrate formation was observed, indicating a transformation of CBZ to dihydrate had occurred instead of CBZ dissolution into the medium. This also explains the slow dissolution observed from the USP experiment. This result contradicts Prajapati, et al.'s experiment (2007) where they demonstrated that PEG 4000 can increase CBZ dissolution rate due to its hydrophilic nature and its ability to increase wettability.

However, Qiu, et al. (2016) proposed that for CBZ-NIC cocrystals with PEG, PEG dissolves quickly to form small channels allowing water to penetrate inside the tablet the moment the tablet is submerged into the dissolution medium. The enhanced dissolution rate contributed further to the dihydrate formation as a micro-environmental supersaturated state was formed before tablet erosion and disintegration (Qiu, et al., 2016). The ATR-FTIR images verified this, showing an immediate CBZ dihydrate formation when dissolution medium was flowed through.

3.2.1.2. Hydroxypropyl methylcellulose (HPMC) excipient

The cocrystal-HPMC tablets visibly showed a slower dissolution compared to the cocrystal-PEG tablets as cocrystal was still present on the images even after an hour of dissolution as observed in Figure 13. and Figure 14. This is due to HPMC swelling and forming a gel layer in the tablet thus forming a viscous environment in the tablet. Subsequently, the diffusion of CBZ into the aqueous medium is hindered and this causes a drug build-up in the matrix of the tablet leading to the crystallisation of CBZ to CBZ dihydrate and hence, some dihydrate formation is observed from the images.

The lesser amount of dihydrate formed compared to the cocrystal-PEG tablets can further be explained by Katzhendler, et al.'s (1998) experiment. They have proven that the HPMC polymer can be a CBZ dihydrate

formation inhibitor during CBZ dissolution, inducing CBZ amorphism. The CBZ water binding sites are hydrogen bonded to the hydroxyl groups in HPMC which prevents the nucleation sites of CBZ dihydrate forming and CBZ-HPMC complexes are formed instead.

However, Li, et al. (2014) showed that dihydrate inhibition was not observed for CBZ-NIC cocrystals. This was because the rate of drug dissolution was greatly enhanced by nicotinamide causing the rate of dissolution of cocrystal to be greater than the rate of HPMC-CBZ complex formation. Thus, supersaturation of CBZ still occured, leading to CBZ dihydrate formation as seen in Figure 13 and Figure 14.

3.2.1.3. Hydroxypropyl Cellulose (HPC) excipient

The FTIR images as shown in Figure 15 and Figure 16 of cocrystal-HPC tablet demonstrated the highest dissolution rate relative to tablets made with other excipients, which agrees with the USP results where tablets made with HPC exhibited the fastest dissolution. The cocrystals disappeared fully from the imaging area in less than an hour. Additionally, no CBZ dihydrate formation was observed, which suggests that cocrystals have actually dissolved quickly. It was deduced that this resulted from the slow hydration properties of HPC, indicating a low water uptake even for a long immersion time in water (Vueba, et al., 2004; Roy & Rohera, 2002). Therefore, cocrystal dissolves from the surface of the tablet without water ingress, and supersaturated conditions for recrystallisation were avoided. This result is comparable to Qiu, et al.'s experiment (2016) where hydroxypropyl methyl cellulose acetate succinate (HPMCAS), a cellulose with slow hydration properties, was used as an excipient for the same cocrystal and in their experiment, no CBZ dihydrate was observed.

3.2.1.4. Saccharin (SAC) excipient

Cocrystal dissolution images were not produced as the cocrystal peak for integration disappeared instantly when dissolution medium was flowed through. Figure 18 shows the spectra obtained at both dry and at time 0 minutes, the start of the dissolution test.

Multiple shifts of the peaks were observed, suggesting a change in cocrystal structure or breakage of the hydrogen bonding which could have resulted from the cocrystal molecules interacting with SAC in the aqueous medium. These peak shift trends are consistent in tablets dissolving in water, pH4 acid and pH5 acid.

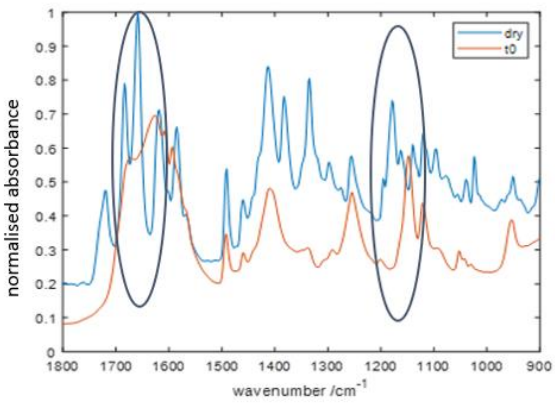


Figure 18 Cocrystal-SAC spectra obtained before dissolution (dry) and beginning of dissolution (t0)

Dihydrate is nevertheless observed as seen in Figure 17. Similar to the cocrystal-PEG tablets, the formation of dihydrate was observed in the ATR-FTIR images almost immediately after the dissolution test started. However, from the USP results, the cocrystal-SAC tablets showed an even slower release of CBZ compared to the cocrystal-PEG tablets.

4. Conclusion

ATR-FTIR imaging technique is a very promising method to observe chemical changes during tablet dissolution. It provides qualitative information inside the tablet and supplements the traditional USP tests, which provides only quantitative results. The amount of dihydrate formation correlates to the dissolution rate of cocrystals. pH was not low enough to show a strong effect on dissolution rate whereas changing the excipient used in the formulation presents a strong effect on the dissolution rate of the cocrystal tablets. HPC is the most promising excipient to allow for a controlled release of CBZ as no dihydrate formation was observed.

5. Limitations

Some things that could be improved in the experiment include the tablet formulation process. Tablets could be made thinner by using less powder to allow for a clearer interface to show on ATR-FTIR images and to promote contact between tablet and surface of diamond crystal. Using a thicker O-ring of the same diameter might help with the contact and clarity of the interface as well. The tablets could be formulated using a different set of equipment with more precise quantification of the compaction force such as using a suitably fitted anvil coupled with a torque screwdriver. Alternatively, perhaps other methods of formulation producing more homogeneous mixtures of cocrystal could be explored further such as the hot melt method

(Karimi-Jafari, et al., 2018). Furthermore, an improvement for in-vitro and in-vivo correlations can be done by following standard USP dissolution procedure at rotational speed 100 rpm and temperature of 37˚C.

References

Cao, H.-L.et al., 2018. Two-Component Pharmaceutical Cocrystals Regulated by Supramolecular Synthons Comprising Primary N···H···O Interactions. *Crystal Growth & Design,* 19(1), pp. 3-16.

Cohen, J. L. et al., 1990. The Development of USP Dissolution and Drug Release Standards. *Pharmaceutical Research,* 7(10), pp. 983-987.

D'Arcy, D. M., Corrigan, O. I. & Healy, A. M., 2006. Evaluation of hydrodynamics in the basket dissolution apparatus using computational fluid dynamics—Dissolution rate implications. *European Journal of Pharmaceutical Sciences,* 27(2-3), pp. 259-267.

Drozd, K. V., Manin, A. N., Churakov, A. V. & Perlovich, G. L., 2017. Novel drug–drug cocrystals of carbamazepine with para-aminosalicylic acid: screening, crystal structures and comparative study of carbamazepine cocrystal formation thermodynamics. *CrystEngComm,* 19(30), pp. 4273-4286.

Drugbank, 2005. *Carbamazepine - Drugbank.* [Online] Available at: https://www.drugbank.ca/drugs/DB00564 [Accessed 11 12 2019].

Eguchi, N., Kawabata, K. & Goto, H., 2017. Electrochemical Polymerization of 4,4-Dimethyl-2,2'-Bithiophene in Concentrated Polymer Liquid Crystal Solution. *Journal of Materials Science and Chemical Engineering,* Volume 5, pp. 64-70.

Emami, S., 2018. Recent advances in improving oral drug bioavailability by cocrystals. *BioImpacts,* 8(4), pp. 305-320.

Ewing, A. V., Wray, P. S., Clarke, G. S. & Kazarian, S. G., 2015. Evaluating drug delivery with salt formation: Drug disproportionation studied in situ by ATR-FTIR imaging and Raman mapping. *Journal of Pharmaceutical and Biomedical Analysis,* 111(1), pp. 248-256.

Fleischman, S. G. et al., 2003. Crystal Engineering of the Composition of Pharmaceutical Phases: Multiple-Component Crystalline Solids Involving Carbamazepine. *Crystal Growth & Design,* 3(6), pp. 909-919.

Fukte, S. R., Wagh, M. P. & Rawat, S., 2014. Coformer Selection: An Important Tool in Cocrystal Formation. *International Journal of Pharmacy and Pharmaceutical Sciences,* 6(7), pp. 9-14.

Gomez-Orellana, I., 2005. Strategies to improve oral drug bioavailability. *Expert Opinion on Drug Delivery,* 2(3), pp. 419-433.

Greco, K. & Bogner, R., 2012. Solution-Mediated Phase Transformation: Significance During Dissolution and Implications for Bioavailability. *Journal of Pharmaceutical Sciences,* 101(9), pp. 2996-3018.

Homayun, B., Lin, X. & Choi, H.-J., 2019. Challenges and Recent Progress in Oral Drug Delivery Systems for Biopharmaceuticals. *Pharmaceutics,* 11(3), p. 129.

Huang, N. & Rodríguez-Hornedo, N., 2010. Effect of Micellar Solubilization on Cocrystal Solubility and Stability. *Crystal Growth & Design,* 10(5), pp. 2050-2053.

Karimi-Jafari, M., Padrela, L., Walker, G. M. & Croker, D. M., 2018. Creating Cocrystals: A Review of Pharmaceutical Cocrystal Preparation Routes and Applications. *Crystal Growth & Design,* 18(10), pp. 6370-6387.

Katzhendler, I., Azoury, R. & Friedman, M., 1998. Crystalline properties of carbamazepine in sustained release hydrophilic matrix tablets based on hydroxypropyl methylcellulose. *Journal of Controlled Release,* 54(1), pp. 69-85.

Kazarian, S. G. & Ewing, A. V., 2013. Applications of Fourier transform infrared spectroscopic imaging to tablet dissolution and drug release. *Expert Opinion on Drug Delivery,* 10(9), pp. 1207-1221.

Kobayashi, Y., Ito, S., Itai, S. & Yamamoto, K., 2000. Physicochemical properties and bioavailability of carbamazepine polymorphs and dihydrate. *International Journal of Pharmaceutics,* 193(2), pp. 137-146.

Lennernäs, H., Shah, V. P., Crison, J. R. & Amidon, G. L., 1995. A Theoretical Basis for a Biopharmaceutic Drug Classification: The Correlation of in Vitro Drug Product Dissolution and in Vivo Bioavailability. *Pharmaceutical Research,* 12(3), pp. 413-420.

Li, M. et al., 2014. Investigation of the Effect of Hydroxypropyl Methylcellulose on the Phase Transformation and Release Profiles of Carbamazepine-Nicotinamide Cocrystal. *Pharmaceutical Research,* 31(9), pp. 2312-2325.

Loftsson, T. & Brewster, M. E., 2010. Pharmaceutical applications of cyclodextrins:basic science and product development. *Journal of Pharmacy and Pharmacology,* 62(11), p. 1607–1621.

Prajapati, S. T., Gohel, M. C. & Patel, L. D., 2007. Studies to enhance dissolution properties of carbamazepine. *Indian Journal of Pharmaceutical Sciences,* 69(3), pp. 427-430.

Punčochová, K. et al., 2017. The Combined Use of Imaging Approaches to Assess Drug Release from Multicomponent Solid Dispersions. *Pharmaceutical Research,* 34(5), p. 990–1001.

Qiao, N. et al., 2013. In situ monitoring of carbamazepine–nicotinamide cocrystal intrinsic

dissolution behaviour. *European Journal of Pharmaceutics and Biopharmaceutics,* 83(3), pp. 415-426.
Qiu, S. et al., 2016. Role of polymers in solution and tablet-based carbamazepine cocrystal formulations. *CrystEngComm,* 18(15), pp. 2664-2678.
Qiu, S. & MingzhongLi, 2015. Effects of coformers on phase transformation and release profiles of carbamazepine cocrystals in hydroxypropyl methylcellulose based matrix tablets. *International Journal of Pharmaceutics,* 479(1), pp. 118-128.
Remenar, J. F. et al., 2007. Celecoxib:Nicotinamide Dissociation: Using Excipients To Capture the Cocrystal's Potential. *Molecular Pharmaceutics,* 4(3), pp. 386-400.
Roy, D. S. & Rohera, B. D., 2002. Comparative evaluation of rate of hydration and matrix erosion of HEC and HPC and study of drug release from their matrices. *European Journal of Pharmaceutical Sciences,* 16(3), pp. 193-199.
Sathisarana, I. & Dalvi, S. V., 2019. Cocrystallization of carbamazepine with amides: Cocrystal and eutectic phases with improved dissolution. *Journal of Molecular Structure,* 1193(1), pp. 398-415.
Sathisaran, I. & Dalvi, S. V., 2018. Engineering Cocrystals of Poorly Water-Soluble Drugs to Enhance Dissolution in Aqueous Medium. *Pharmaceutics,* 10(3), p. 108.
Schultheiss, N. & Newman, A., 2009. Pharmaceutical Cocrystals and Their Physicochemical Properties. *Crystal Growth & Design,* 9(6), pp. 2950-2967.
Shewale, S. et al., 2015. Formulation and Solid State Characterization of Nicotinamide-based Co-crystals of Fenofibrate. *Indian Journal of Pharmaceutical Sciences,* 77(3), p. 328.
Sugano, K. & Terada, K., 2015. Rate- and Extent-Limiting Factors of Oral Drug Absorption: Theory and Applications. *Journal of Pharmaceutical Sciences,* 104(9), pp. 2777-2788.
The United States Pharmacopeial Convention, 2011. *<711 Dissolution>.* s.l.:The United States Pharmacopeial Convention.

U.S. Department of Health and Human Services FDA, 2018. *Regulatory Classification of Pharmaceutical Co-Crystals Guidance for Industry.* s.l.:U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER).
U.S. National Library of Medicine, 1984. *HSDB: 1237 - PubChem.* [Online] Available at: https://pubchem.ncbi.nlm.nih.gov/source/hsdb/1237 [Accessed 11 12 2019].
Vueba, M. et al., 2004. Influence of cellulose ether polymers on ketoprofen release from hydrophilic matrix tablets. *European Journal of Pharmaceutics and Biopharmaceutics,* 58(1), pp. 51-59.
Wang, L., Tan, B., Zhang, H. & Deng, Z., 2013. Pharmaceutical Cocrystals of Diflunisal with Nicotinamide or Isonicotinamide. *Organic Process Research & Development,* 17(11), pp. 1413-1418.
Yuliandra, Y. et al., 2018. Cocrystal of Ibuprofen–Nicotinamide: Solid-State Characterization and In Vivo Analgesic Activity Evaluation. *Scientia Pharmaceutica,* 86(2), p. 23.
Zhang, Q., Gladden, L., Avalle, P. & Mantle, M., 2011. In vitro quantitative 1H and 19F nuclear magnetic resonance spectroscopy and imaging studies of fluvastatin™ in Lescol® XL tablets in a USP-IV dissolution cell. *Journal of Controlled Release,* 156(3), pp. 345-354.

The Development of a High-Pressure Micro-Viscometer

Riley Latcham
Department of Chemical Engineering, Imperial College London, U.K.

Abstract

A high-pressure vibrating wire viscometer, of working volume <100 μL, has been developed and validated against fluids spanning 6 orders of magnitude of viscosity and at pressures up to 10 MPa. The instrument featured a novel internal tensioner, which maintained wire tension through the use of a compressed spring. The unexpected response from the length of wire within the IT led to the adaptation of the working equation to allow multiple resonance modes. The working equation was further adapted upon the discovery of the need for a previously unaccounted for impedance term. This new equation should allow for the more accurate measurement of high viscosity liquids, as well as allowing potential new variants of vibrating wire viscometer to be made.

Introduction

Within the field of fluid metrology, VWVs have, since their inception in 1901 [1], been promising instruments for use in both industry and academia, stemming from their rigorous working equation with detailed corrections and their simple, rugged design which makes them well suited to HTHP as well as allowing for a small working volume [2].

Their origin as a class of instruments can be traced back to Stokes solution of the Navier-Stokes equation for an infinitely long cylindrical rod vibrating in an unbounded fluid [1], which formed the fluid dynamic basis of the operating principle. While this was sufficient to produce a proof of concept, significant technological developments were required before the electronics needed for high accuracy measurement became available. In 1963 the first practical device was constructed and used to measure the viscosity of liquid helium [3], combining the hydrodynamic theory with the mechanical theory of a wire without flexural stiffness [4]. The equations were then adapted for a prismatic rod with flexural stiffness, a more realistic representation, by Retsina *et al.* in 1986/7 [5] who also defined clear operating limits and added a corrective term for the case of a bounded fluid.

From this point onwards VWVs began to diverge both in their design and their operating principle: some being operated in transient modes while others were steady state; while traditionally featuring a wire clamped at both ends some added a sinker to tension the wire. The latter was found to allow for simultaneous measurement of both viscosity and density [6] at the cost of orientation independent operation. Other, less notable variants have been reported, such as an attempted spring-tensioned VWV [7] and one that featured two wires [8], though little is known about it and the concept appears abandoned, with one very brief mention of the idea recently [9].

Such diversity, bolstered by the complexity of the underlying theory, alongside simplifications and the occasional error, has led to a range of different working equations and a non-insignificant number of disagreements with regards to specific expressions. It is worth noting here that while the current working equations are rigorous, they are not strictly complete. At present there exists potential shortcomings in the theory surrounding operation with wires in close proximity to boundaries [2], and the measurement viscosities above $100\ mPa \cdot s$.

Despite this VWVs are undeniably highly capable instruments, moreover they have been demonstrated to be absolute instruments with a measurement of the viscosity of water being reported with relative error 0.03% at $T = 293.15\ K$ and $p = 101.325\ kPa$ [10]. Practical limitations, such as difficulty in measuring the wire radius to sufficient accuracy, have led to most VWVs being operated in a quasi-absolute manner instead. This is not to say that they are relative instruments, calibration is merely used to more accurately determine a few key parameters than is otherwise practical.

In spite of the clear potential for VWVs; no commercial version of the instrument has been produced so far [11], in part due to a lack of proven long-term reliability occurring as a result of VWVs' sensitivity to their initial calibration parameters. As such it would be pertinent to develop a VWV that is fundamentally more reliable.

Based on the current state of the literature and applications to which vibrating wire viscometry is relevant, this work set about developing a new high-pressure micro-viscometer. A prototype micro-scale VWV had existed in the laboratory as the result of an earlier project, the modification and assembly of which formed the foundation of the work. The objective was to ensure the new VWV to be accurate and capable of maintaining wire tension, as the prototype encountered problems with this although it produced promising results regardless. It was operated at pressures up to $10\ MPa$ and temperatures up to $350\ K$, and was capable of using less than $1\ mL$ sample fluid. While for certain tests the total system volume exceeded this, the internal volume of the viscometer cell itself was less than $100\ \mu L$. The instrument was calibrated using gaseous nitrogen and ambient pressure validation was performed using calibration standard fluids of viscosities ranging from around 7 to $1230\ mPa \cdot s$, high-pressure validation was then performed using squalane, though various issues were encountered with these measurements.

The tension of the wire was successfully maintained through the use of an internal tensioner utilising a compressed spring, however the section of wire housed within it was found to produce unexpected resonance peaks which, for this specific design,

negatively affected measurements. Measurements of the highly viscous calibration fluid S600 yielded poor results, however detailed analysis of unveiled shortcomings in the underlying theory.

The working equation of the vibrating wire viscometer was modified to account for the unexpected results, allowing for use not only with this specific VWV but for the class of instruments in general, and should now allow for the more accurate measurement of high viscosity fluids. Following from this, two novel instruments have been proposed: a dual-chambered vibrating wire viscometer and the other a multi-wire viscometer-densimeter.

Working Equations

There are three aspects to the theory and operating principle of a vibrating wire viscometer. Two of these, the fluid dynamics and wire mechanics, address the motion of the wire itself; the signal processing addresses how the information is gathered and utilised. Due to the length and complexity of the full theory and its derivations, only the working equations pertinent to the operation of the viscometer will be given, for a more complete description refer to the literature [12] [5] [13] [14]. The main working equation for the VWV is:

$$V = V_1 + V_2 \tag{1}$$

Where V is the voltage across the wire, consisting of the voltage produced by wire motion V_1 and the background voltage developed V_2.

$$V_1 = \frac{\Lambda f i}{f_0^2 - f^2(1+\beta) + f^2(\beta' + 2\Delta_0)i} \tag{2}$$

$$V_2 = a + bi + cfi \tag{3}$$

Where Λ is the signal amplitude, f the resonant frequency of the wire in the fluid, f_0 the resonant frequency *in vacuo*, Δ_0 the logarithmic decrement of the wire *in vacuo*, a, b, and c account for voltage offsets from the lock-in amplifier and electrical impedance, and finally β and β' are the dimensionless added mass and dimensionless viscous damping due to the fluid surrounding the wire, respectively.

When accounting for the presence of the cell wall β and β' are defined as:

$$\beta = \frac{\rho}{\rho_s}\Re(H-1) \tag{4}$$

$$\beta' = -\frac{\rho}{\rho_s}\Im(H) \tag{5}$$

Where ρ is the fluid density, ρ_s the wire density, $\Re$ refers to the real components, $\Im$ to the imaginary components, and H and its constituents are defined as follows:

$$H = \frac{A+B}{C+D+E} \tag{6}$$

$$A = 2h^2\big(I_0(h)K_0(h^*) - I_0(h^*)K_0(h)\big) \ldots \\ +4h\big(I_1(h)K_0(h^*) + I_0(h^*)K_1(h)\big) \tag{7}$$

$$B = \left(\frac{4h}{\sigma^*}\right)\big(I_0(h)K_1(h^*) - I_1(h^*)K_0(h)\big) \ldots \\ -\left(\frac{8}{\sigma^*}\right)\big(I_1(h)K_1(h^*) + I_1(h^*)K_1(h)\big) \tag{8}$$

$$C = \frac{h^2}{\sigma^{*2}}(\sigma^{*2}-1)\big(I_0(h)K_0(h^*) - I_0(h^*)K_0(h)\big) \tag{9}$$

$$D = \left(\frac{2h}{\sigma^*}\right)\begin{pmatrix} I_0(h)K_1(h^*) - I_1(h^*)K_0(h^*) \ldots \\ +I_1(h^*)K_0(h) + I_0(h^*)K_1(h^*) \end{pmatrix} \tag{10}$$

$$E = \left(\frac{2h}{\sigma^{*2}}\right)\begin{pmatrix} I_0(h^*)K_1(h) - I_0(h)K_1(h) \ldots \\ +I_1(h)K_0(h^*) + I_1(h)K_0(h) \end{pmatrix} \tag{11}$$

$$h = \left(\frac{\rho\omega R^2}{\eta} i\right)^{1/2} \tag{12}$$

$$h^* = \sigma^*\left(\frac{\rho\omega R^2}{\eta} i\right)^{1/2} \tag{13}$$

$$\sigma^* = \frac{R_C}{R} \tag{14}$$

Where I_0, I_1, K_0, and K_1 refer to zeroth and first order modified Bessel functions of the first and second kinds respectively [15], ω is the angular frequency, R the wire radius, and R_C the internal radius of the viscometer cell.

Apparatus

The apparatus consisted of three main parts: the viscometer cell, the auxiliary cell, and peripherals. The majority of the components in the viscometer cell and auxiliary cell were manufactured in-house.

A cold-drawn gold-plated tungsten wire, obtained from Luma Metall, was used as the high density ($\rho \approx 19250\ kg \cdot m^{-3}$) and tensile strength allowed for high quality resonance and good measurement sensitivity through the term ρ/ρ_s. A nominal wire diameter of $0.2\ mm$ was used to allow for high viscosity fluids to be measured, VWVs of this scale normally use smaller wire [2].

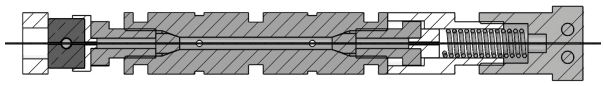


Figure 1. Engineering drawing of the viscometer cell showing the fluid chamber in the middle and the internal tensioner on the right

The viscometer cell, shown in Figure 1, had a main body constructed out of the titanium alloy Ti-6Al-4V, chosen for its strength and corrosion resistance, with three FKM O-rings around it (not shown) used to seal and electrically insulate it against the auxiliary cell. It had an internal diameter of $2\ mm$, resulting in the ratio of the radius of the outer boundary to the wire radius of only 10. This is below the stipulated lower limit [5] however good results have been obtained with a ratio as low as 6.6 [2] and it allowed for a smaller fluid volume so this was considered acceptable. The internal volume was approximately $80\ \mu L$, a suitably small volume for use in core sample fluid testing and with precious fluids.

The total working volume of the apparatus was not measured as it is of lesser importance and varied depending on what was being tested.

The wire was gripped on either end by polymeric Vespel ferrules, chosen for their high maximum temperature and ability to grip the wire whilst the cell was being assembled. Such ferrules should be annealed [] which will likely minimise changes to the system, however an investigation around this was beyond the scope of this work and thus the ferrules were used as purchased. The ferrules were compressed by stainless steel SS316 nuts, sealing the cell.

Wire tension was initially achieved by clamping one end of the wire in a pin grip whilst the other supported a tensioning mass. This was left for 3 hours before switching tension control over to a novel internal tensioner such that the cell could then be used without risk of tension loss due to the wire slipping in the polymeric ferrules.

The internal tensioner consisted of a PEEK sleeve and spacer fitted over one of the stainless steel nuts, originally capable of freely rotating so as to prevent accidental twisting of the wire, however this design was later modified to one that fit securely over the hexagonal nut so that the whole cell could be rotated. A stainless steel clamping structure was threaded over the sleeve, compressing a stainless steel spring housed within, and a clamping plate then used to firmly clamp the wire in place.

Wire tension could then be maintained regardless of wire slip through the polymeric ferrules; however, it was found that the tension between the main chamber and the internal tensioner may take some time to equilibrate following initial assembly. The capacity of the Vespel ferrules to hold the tension in the wire, originally thought to be a benefit, appears to have hindered the effectiveness of the internal tensioner and the reliability of the calibrated parameters.

Finally, tension was held on the other side of the cell using a grub screw. The assembled viscometer cell could then be placed into the auxiliary cell, with a PEEK spacer used to ensure the correct positioning of the fluid ports.

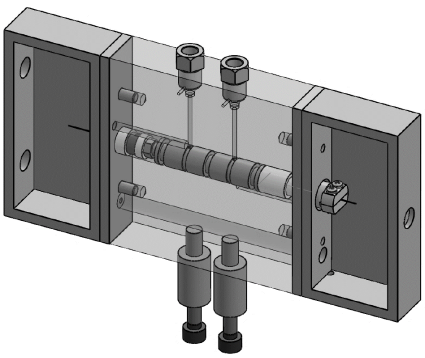


Figure 2. CAD model of the auxiliary cell with the viscometer cell inserted. Fluid ports can be seen, as well as holes for the thermocouples on opposite sides of the cell

The auxiliary cell, as shown above, housed the viscometer cell and consisted primarily of a stainless steel external pressure cell and two aluminium electrical terminal boxes, separated by PTFE pieces. The auxiliary cell allowed the viscometer to be placed between the poles of a SmCo permanent magnet, allowing for a magnetic flux density of roughly $0.6\ T$ around the wire. This is higher than the more common choice of $0.3\ T$ [] but is well suited to such a small, thick-wired viscometer.

The electrical terminal boxes housed the electrical connections and shielded them external electrical interference, as well as allowing two K-type thermocouples to be placed into the stainless steel block. One fed data to an Agilent 34970A data acquisition unit whilst the other was used by the RKC MA900 temperature controller, which heated the system using four flexible heater-pads adhered to the external pressure cell, insulated using a silicone sponge sheet. One such pad failed irreparably, resulting in issues discussed in later. Pressure was measured using a DJ Instruments zero-dead-volume thru-Flow sensor connected inline and read by the data acquisition unit.

The induced voltages were measured by a Stanford Research Systems SR830 DSP lock-in amplifier and fed to a computer for processing through a GPIB interface. The lock-in amplifier also provided the sinusoidal excitation current to drive the wire.

Methodology

A series of tests were performed in this work, both to calibrate and validate the VWV and to investigate unexpected results. The following fluids were used, as purchased, during this work:

- Ambient air was used to ensure that the instrument was operating and to provide an early estimate of Δ_0 as data is readily available [16].
- Hexane, obtained from Sigma-Aldrich, was used to flush the system between measurements.
- The viscosity calibration standards S6, S60, and S600 were obtained from Paragon Scientific Ltd, whereas S20 was obtained from Cannon Instrument Company. The dynamic and kinematic viscosities as well as the densities of these fluids at specific temperatures were supplied alongside these, these were the values used (or interpolated from where needed) when testing, as looking up values online has a high risk of returning values for a different batch. These were used for initial calibration and validation.
- Nitrogen was fed into the system from a cylinder of $200\ bar$ zero-grade nitrogen obtained from BOC Ltd. It was used to determine Δ_0 as a function of temperature and to pressurise squalane.
- Squalane, obtained from Sigma-Aldrich, was used for validation tests at high pressures, as there exist detailed correlations for calculating its viscosity and density supported by numerous studies [].

To take measurements the wire was driven in forced-mode steady-state transverse oscillation by a constant sinusoidal current produced by the lock-in amplifier, generating a transverse Lorentzian force across the wire and causing it to move. This motion induces additional voltage across the wire in accordance with Faraday's Law of induction which can be picked up and measured. Thus, by performing a frequency

sweep, an induced voltage-frequency curve can be generated which is linked to viscosity through the working equations.

The initial signal processing is also performed by the lock-in amplifier, which measures the in-phase and quadrature voltage across the wire and applies a technique known phase-sensitive detection to output a sufficiently narrow bandwidth for analysis. This is done by amplifying the input signal and mixing it with an appropriate reference frequency before demodulating and filtering to output DC voltage, allowing for high amplification of the small generated signal with little noise [17]. This output is then fed to a computer for further processing and determination of viscosity.

Results and Discussion

Initial testing of the fully-assembled viscometer made it evident that equipment limitations would be an issue. The two thermocouples were measuring temperatures in disagreement by $\sim 0.5 - 6\,K$, increasing with temperature. A single point calibration at $273.15\,K$ using ice water failed to rectify the issue. The main source of this disagreement was the failure of one of the four heater pads. This resulted in a temperature gradient across the cell and prevented accurate measurement, as even use of viscosity calibration standards required an accurate knowledge of the temperature. The pressure sensor proved to also be somewhat unreliable, with readings varying by around $0.01\,MPa$. Whilst normally such problems could be solved by replacing faulty parts or carrying out further calibrations, circumstances surrounding the work prevented this. To partially account for this, temperatures have been given to the nearest Kelvin, and are based on an examination of results for fluids of known viscosity.

Initial results produced by the VWV revealed that the internal tensioner was indeed functioning, with a resonant frequency of making it clear that the wire was under tension through comparison to the theoretical model for a stiff-tensioned wire [13].

The single reference point calibration in ambient air, $T = 293.15\,K$ and $p = 101.325\,kPa$, returned a vacuum decrement $\Delta_0 = 0.00249$. The single reference point calibration using S60 under the same conditions to more accurately determine the radius returned a value of $r = \sim 95\,\mu m$, however it became apparent that the measurements were somewhat unreliable. In the days following this Δ_0 and f_0 slowly increased, whilst the Q factor decreased. It is believed that this was caused by a tension imbalance between the sections of wire within the main chamber and the internal tensioner, with the tension in the latter being higher. The results of validation tests within this period of time can be seen below in Table 1. All things considered these results were encouraging; however, they also demonstrate the extent to which the VWV was limited in its accuracy.

Table 1. Results of initial validation tests performed using viscosity calibration fluids, illustrating that reasonably good measurements could still be made under the non-ideal circumstances

Fluid	T/°C	$\eta_{exp}/mPa \cdot s$	$\eta_{real}\ \backslash mPa \cdot s$	% error
S6	25	6.754	7.337	7.952
S20	25	27.95	28.83	3.067
S60	25	105.8	102.5	3.211
S600	50	205.1	251.3	18.40

Table 2. Results of high pressure squalane tests

T/°C	*P/ MPa*	Nexp/ *mPa • s*	Nreal	
25	1.80	0.02884	28.8549301929356	0.031
	3.90	0.0429	30.2691083669326	42.0
	5.98	N/A	N/A	
	8.10	0.04780	33.1967668904592	44.0
	10.01	0.05071	34.7078655991671	46.1
50	1.80	0.008931	10.2294801206301	12.6
	3.90	0.01030	10.6462738840893	3.24
	5.98	0.01242	11.0689255469341	12.2
	8.10	0.01324	11.5034480517518	15.1
	10.01	0.01194	11.9433246923412	0.049
75	1.80	0.008265	10.2294801206301	19.1
	3.90	0.003696	5.12989238021695	27.9
	5.98	0.003926	5.30832655626722	26.0
	8.10	0.004528	5.48646995877835	17.4
	10.01	0.00429603235628357	5.66764610521217	24.2

The voltage-frequency curves produced by the VWV showed varying number of additional resonance peaks, as can be seen in Figure 3. The identification of these peaks was done by examination of engineering drawings, calculations of theoretical resonant frequencies for the various free lengths of wire present within the instrument, analysis of the effects of varying dwell time, and by observing the effects of adding viscous liquids to the main chamber of the viscometer cell. The last of these clearly demonstrated that the peaks must have been caused by wire in ambient air, as they remained unaffected by the addition of viscous liquids. The section of wire within the internal tensioner was coincidentally of very similar ($< 3\,mm$ shorter) length to that of the main chamber but was further from the magnets and thus in an area of notably lower magnetic flux. This explains why the peaks were occurring at a frequency several hundred Hertz above the main peak and were of significantly lower amplitude when the viscometer cell was in ambient air.

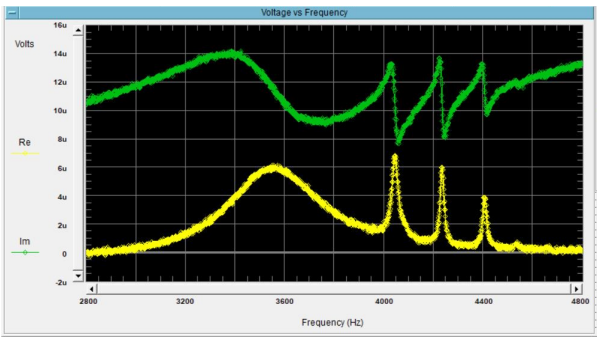


Figure 3. The resonance curves for S60, showing the presence of several undamped peaks along with the damped main peak, with the in-phase voltage in yellow and the quadrature voltage in green.

To account for the presence of multiple peaks the equation for voltage produced by the motion of the wire can be written as:

$$\sum_{j=1}^{n} V_j = \sum_{j=1}^{n} \frac{\Lambda_j f i}{f_{0,j}^2 - f^2(1+\beta_j) + f^2(\beta_j' + 2\Delta_{0,j})i} \quad (15)$$

Where V_j is the voltage produced by the wire motion in the jth mode and n is the number of resonance modes. It can be seen that for $n = 1$ this equation reduces to the original version of the equation.

It was observed, and expected, that highly viscous fluids damp the motion of the vibrating wire significantly resulting in low Q factors. This means that a higher drive voltage is required to measure a signal of sufficient amplitude []. Even at the maximum drive voltage of the LIA used in this work ($5\ V$) this damping resulted in a relatively flat peak for the calibration fluid S600, however this highlighted an interesting effect: the in-phase voltage following the resonance peak was higher than expected by the theory. To investigate this further a series of large frequency sweeps were performed far from the influence of any resonance peaks. The in-phase voltage was found to increase near constantly at a rate of $0.165\ nV \cdot Hz^{-1}$ and thus it was clear the theory needed to be re-examined.

The current equation for the background voltage V_0 is an approximation [7], where the term cfi accounts for the electrical impedance of the wire. The need for such a term is the result of the wire effectively behaving like a resistor and an inductor in series, and it is here that the issue becomes evident. For such a system the electrical impedance Z can be defined as:

$$Z = R + Xi \tag{17}$$

Where R is the resistance and X is the reactance, which in this case is an inductance. Impedance depends on the frequency of the sinusoidal voltage; however, it appears this has been omitted for the in-phase voltage thus far.

The resistance of an AC circuit is a function of frequency as a result of the skin effect, which reduces the effective cross-sectional area of the wire and thus increases the resistance. It can be seen from equation X that this would result in the observed in-phase voltage increase. From this, the equation for the background voltage V_0 has been rewritten to account for the in-phase voltage impedance c and to help clarify the meanings of the respective terms, such that:

$$V_0 = a + bi + (c + di)f \tag{18}$$

Where a and b account for the offsets of the lock-in amplifier for the in-phase and quadrature voltages respectively; c and d account for the electrical impedance of the wire, which is also a function of the frequency f. The total voltage developed across the wire is thus given by:

$$V = V_0 + \sum_{j=1}^{n} V_j \tag{19}$$

As a demonstration of the need to account for this impedance, the in-phase part of the resonance curve for squalane at $T = 298.15\ K$ and $P = 10\ MPa$ was simulated with and without the additional term, using the real parameters for this work and is shown below.

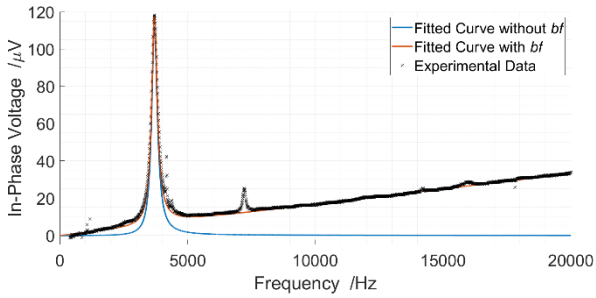


Figure 4. The in-phase voltage for squalane at $T = 298.15\ K$ and $P = 10\ MPa$

The impact of the term c increases with increasing viscosity due to the increased bandwidth, which provides some explanation to the poor performance of VWVs when measuring high viscosity fluids. When one attempts to fit experimental data to the original working equation convergence is still reasonably feasible, however the calculated viscosity is skewed.

Conclusions

In this work a micro-scale vibrating wire viscometer capable of high-pressure measurements and featuring a novel internal tensioner has been developed, albeit affected by issues with some of the peripherals. The previously unreported effect of electrical impedance on the in-phase voltage has been analysed, and the working equation has been modified to account for this, alongside a generalisation to allow for the presence of multiple peaks. The new equation allows for higher accuracy measurement of high viscosity fluids, and potentially even new VWV variants which intentionally induce addition peaks.

Acknowledgements

I would like to thank Gavin Barnes for his role in machining parts for this viscometer, as well as Deepali Dholakia, who's constant support is what made this work possible even in the face of insurmountable odds.

References

[1] G. G. Stokes, Mathematical and Physical Papers, Cambridge: Cambridge University Press, 1901.

[2] G. Dehestru, et al., "A microfluidic vibrating wire viscometer for operation at high pressure and high temperature," Review of Scientific Instruments, vol. 82, no. 3, 2011.

[3] J. T. Tough, W. D. Mc Cormick, and J. G. Dash, "Viscosity of liquid He 2," Physical Review, vol. 132, no. 6, pp. 2373-2378, 1963.

[4] J. T. Tough, W. D. Mc Cormick, and J. G. Dash, "Vibrating Wire Viscometer," Review of Scientific Instruments, vol. 35, no. 10, pp. 1345-1348, 1964.

[5] Retsina. T, Richardson. S. M and Wakeham. W. A, “The theory of a vibrating-rod viscometer,” Applied Scientific Research, vol. 43, no. 4, pp. 325-346, 1987.

[6] A. A. H. Padua et al., “Validation of an accurate vibrating-wire densimeter: Density and viscosity of liquids over wide ranges of temperature and pressure,” International Journal of Thermophysics, vol. 17, no. 4, pp. 781-802, 1996.

[7] M. McBride-Wright, Viscosity and Density of Aqueous Fluids with Dissolved CO2, London: Imperial College London, 2013.

[8] J. Wilhelm et al., “A Vibrating-Wire Viscometer for Dilute and Dense Gases,” International Journal of Thermophysics, vol. 19, no. 2, pp. 391-401, 1998.

[9] C. C. Hopkins, Vibrating-wire Rheometry, University of Western Ontario, 2018.

[10] A. R. H. Goodwin and K. N. Marsh, “An Absolute Viscometer for Liquids: Measurement of the Viscosity of Water at T = 298.15 K and p = 0.1 MPa,” Journal of Chemical & Engineering Data, vol. 56, no. 1, pp. 167-170, 2010.

[11] M. J. e. a. Assael, Experimental Thermodynamics Volume IX: Advances in Transport Properties of Fluids, Cambridge: Royal Society of Chemistry, 2014.

[12] T. Retsina, S. M. Richardson and W. A. Wakeham, “The theory of a vibrating-rod densimeter,” Applied Scientific Research, vol. 43, no. 2, pp. 127-158, 1986.

[13] F. Ciotta and J. P. M. Trusler, “Improved understanding of vibrating-wire viscometer-densimeters,” Journal of Chemical and Engineering Data, vol. 55, no. 6, pp. 2195-2201, 2010.

[14] M. J. Assael and S. K. Mylona, “A Novel Vibrating-Wire Viscometer for High-Viscosity Liquids at Moderate Pressures,” Journal of Chemical & Engineering Data, vol. 58, pp. 993-1000, 2013.

[15] M. Abramowitz and I. A. Stegun, Handbook of Mathematic Functions with Formulas, Graphs, and Mathematical Tables, Washington D.C.: National Bureau of Standards, 1964.

[16] G. W. C. Kaye and T. H. Laby, Tables of physical and chemical constants and some mathematical functions, Essex: Longman Scientific and Technical, 1986.

[17] Stanford Research Systems, MODEL SR830 DSP Lock-In Amplifier, Harpenden: Lambda Photometrics Limited, 2009.

Machine Learning to Develop an Engineering Correlation for Critical Properties of Fluids

Adam Ward, Max Bird

***Abstract*—The use of predictive methods to estimate critical properties of compounds is popular in many areas of research, mainly through the use of group contribution methods and Equations of State (EoS). Here, we present a data-driven machine learning approach, using an artificial neural network (ANN) to correlate known molecular descriptors with a set of critical properties for a range of compounds. Molecular descriptors from the SAFT-VR Mie EoS, namely segment length scale, chain length and repulsive constant, are used as training features to predict critical temperature, critical pressure and critical compressibility factor. A large experimental database of over 5,000 organic compounds is used to compare the predictive accuracy of the proposed ANN and the existing SAFT-VR Mie EoS (SAFT). The ANN performs well, able to simultaneously predict all three critical properties both more quickly, and to a higher degree of accuracy compared to SAFT. Results show the ANN achieves a mean absolute percentage error (MAPE) of 7.56% and 4.36% for critical pressure and compressibility factor respectively, compared to 16.9% and 24.2% obtained by SAFT. A small extended study also evaluates the ANN's ability to generalise, by predicting critical properties for inorganic compounds, when trained solely on organic examples. Overall, this paper is a promising first step in realising a fully predictive tool to calculate critical properties of unknown compounds, and provides clear motivation for increased work in this area.**

***Index Terms*—Critical properties, Machine learning, Artificial neural network, Equations of State, SAFT**

I. Introduction

EXPERIMENTAL measurements of the critical properties of fluids can be very challenging to carry out, due to the extreme conditions required and the instability of fluids at and around these conditions [1]. As such, it would be inconvenient, if not impossible, to conceive tractable workflows which depend on on-the-fly measurements of critical points as inputs to physical property models.

The most widely employed approach to circumvent the need for experimental determinations of critical properties is a predictive group contribution method. Nannoolal et al. [1] report a summary of such approaches for the prediction of critical points. The group contribution method is based on the principle that a large number of molecules of interest can be seen simply as a combination of a far smaller number of common "groups". Simple group contribution methods are of first order and do not encode information regarding the structure of molecules, but rather only their composition. More sophisticated second order group contribution methods create super-groups which are combinations of multiple first order groups. Further, it is also possible to increase the accuracy of these methods by providing easily accessible properties of the molecule, such as the normal boiling point, in the contribution calculations. This has the advantage of encoding some knowledge about the real behaviour of the molecule to guide the prediction. Physical properties are calculated as sums of the group contributions present within a given molecule. Group contributions themselves are quantified by fitting to experimental data, which presents a problem when predicting critical properties as the available data set is limited in size. Therefore, the extrapolative capability of such a method is questionable as the formulations of the combination rules are empirical and do not necessarily capture the underlying physical behaviour of the molecules.

Continuing advances in readily available computing power are facilitating detailed insight from theoretical EoS approaches for the prediction of critical properties. One such approach is the coarse graining framework described by Statistical Associating Fluid Theory (SAFT). The formulation of SAFT-VR Mie will be of particular interest in our study, in which molecules are represented as chains of tangentially adjoined spheres with identical length and energy scales, as well as Mie Potential shape factor constants [2]. This framework allows relatively accurate predictions, informed by underlying physics, without increasing the computational cost too much relative to a group contribution method. Typically, model parameters are obtained by fitting to experimental data for readily accessible molecular properties. This information can then be used to inform an iterative solution process to predict the critical properties of a given compound. However, SAFT-type approaches have a major limitation for the prediction of critical properties - they are unable to accurately capture all 3 of the critical temperature, pressure and compressibility factor simultaneously [3].

As an alternative to semi-empirical or theoretical approaches, machine learning (ML) presents an alternative path to obtain critical properties of fluids. The term "ML" describes a broad range of algorithms in which models learn from available data to identify a complex relationship between input features and output variables. Various attempts have been made in the literature to predict properties of fluids using ML algorithms. For example, Ahmadi et al. [4] predicted transport properties (viscosity and thermal conductivity) of carbon dioxide using an artificial neural network (ANN), with temperature and pressure as input features. Their modelling was very accurate on the testing data available, but this work was restricted to properties of a single compound. Other work, such as that by Shadravan et al. [5], has investigated the use of ML algorithms (ANN and Gaussian Process Regression (GPR)) to produce a single unified model to predict properties of a variety of fluids. In this case the aim was to predict rheological properties of fluid mixtures for applications in drilling, using mixture composition and molecular properties, eg. density, as input

features. This study highlights the capability of ML algorithms to learn not only the behaviour of a specific compound in response to a certain set of thermodynamic state variables, but also to learn underlying physical behaviour which finds wider applicability to entire families of molecules.

The use of ML models to predict critical properties of fluids has also been studied. In work this year by Groven et al. [6], an ANN was used moments of inertia, molecular weight and number of aromatic rings as input features to simultaneously predict boiling and critical temperatures, as well as the vapour pressure curve for polycyclic aromatic hydrocarbons. Their work showed that even a simple ANN with only 5 hidden units in a single hidden layer, was able to extrapolate excellently from the simulated training set to accurately predict properties of real fluids. This clearly highlights the power for ML algorithms to learn the fundamental physical relationships which relate appropriate input features to a set of physical properties of interest, motivating us to seek a workflow in which we may predict the full set of critical properties for a fluid using a single model.

In this work, we present an ANN, able to predict the critical temperature, critical pressure and compressibility factor for a range of compounds using a set of available molecular descriptors as training features. These molecular descriptors are calculated via the 4-parameter principle of corresponding states (4PCS) correlation [7], used in the SAFT-VR Mie EoS (SAFT) to characterise the free energy of a given compound. If successful, this work circumvents the need for lengthy iterative calculations using Equations of State and provides a vital first step in the development of a truly predictive model for critical properties of unknown compounds.

II. Modelling Approaches

A. Statistical Associating Fluid Theory

In this study, we use the SAFT-VR Mie EoS [2] as a reference model to compare with the ANN. In this model, the Helmholtz free energy of molecules formed from chains of identical tangentially adjoined spherical segments, with associative interactions, is expressed as:

$$\frac{A^{\mathrm{SAFT}}}{Nk_BT} = \frac{A^{\mathrm{ideal}}}{Nk_BT} + \frac{A^{\mathrm{mono}}}{Nk_BT} + \frac{A^{\mathrm{chain}}}{Nk_BT} + \frac{A^{\mathrm{assoc}}}{Nk_BT} \quad (1)$$

Where A^{mono}, A^{chain} and A^{assoc} are the contributions to the residual free energy from single spherical segments, chains of spherical segments and associative interactions (eg. hydrogen bonding), respectively. The free energy terms are described by parameters which may be obtained by fitting the EoS to readily accessible thermodynamic properties, such as the vapour pressure and saturated liquid density. From the free energy, A^{SAFT}, PVT properties may be obtained through standard thermodynamic relations.

SAFT-VR Mie is a generalisation of the SAFT-VR approach to incorporate the Mie potential, given by:

$$\phi_{\mathrm{Mie}}(\lambda_r, \lambda_a, \epsilon, \sigma, r) = C\epsilon \left[\left(\frac{\sigma}{r}\right)^{\lambda_r} - \left(\frac{\sigma}{r}\right)^{\lambda_a} \right] \quad (2)$$

With:

$$C = \frac{\lambda_r}{\lambda_r - \lambda_a} \left(\frac{\lambda_r}{\lambda_a}\right)^{\lambda_a/(\lambda_r - \lambda_a)}$$

Where λ_r and λ_a are the repulsive and attractive constants respectively, with ϵ and σ being the energy and length scales of the spherical segments.

In this work, we are interested in using SAFT-VR Mie to predict the critical properties of fluids. To obtain the critical temperature, pressure and volume of a fluid for a given set of Mie parameters, an open-source Python implementation of SAFT-VR Mie, available on GitHub [8], was used. In this implementation, the critical point is found using a root-finding algorithm to identify the temperature at which the maximum slope of the isotherm in P-v space is zero. This effectively resolves the critical temperature as $\max(dP/dv)$ is positive in the sub-critical region and negative in the super-critical region. Then, the critical pressure and compressibility factor are found by setting the first and second derivatives of the pressure with respect to volume equal to zero - as the critical point is a point of inflection in P-v space. It should be noted that this implementation of SAFT-VR Mie requires that the user provides an initial guess for the critical volume, as well as a permissible range of values for the critical volume. If an initial guess is provided which is far from the true value, then the solution algorithm may converge to anomalous results. It is also worth noting at this stage that the specific implementation of SAFT-VR Mie used in this work does not support the modelling of associative interactions between molecules, and so cannot accurately model compounds exhibiting hydrogen bonding.

B. Artificial Neural Network

An artificial neural network (ANN) is a predictive algorithm inspired by the structure of neurons in a human brain. Each node takes a set of inputs, $\mathbf{x}$, performs some weighted sum, $\mathbf{w^T x}$, of those inputs and applies some "activation" function, f, on that sum to produce an output. Several nodes are arranged into a layer, with the entire ANN being made up of several of these layers. All outputs of one layer are fed as inputs to all nodes on the following layer. Each layer will also add a linear "bias" term, b, to the input of each node. We are interested in a regression problem, so the physical properties of interest will each be associated with the value of a single output node. This overall network structure is summarised in Fig. 1 and the inner working of a single hidden node is summarised in Fig. 2.

There are many choices for the activation function which nodes can use to produce an output. The most common of these functions is the rectified linear unit function (relu), which is defined as:

$$\mathrm{relu}(x) = \max(0, x) \quad (3)$$

One of the key reasons the relu function is so widely used is that it has a lower likelihood of vanishing gradient, as compared to other non-linear activation functions such as the Sigmoid function or the hyperbolic tangent function, which can lead to faster model training.

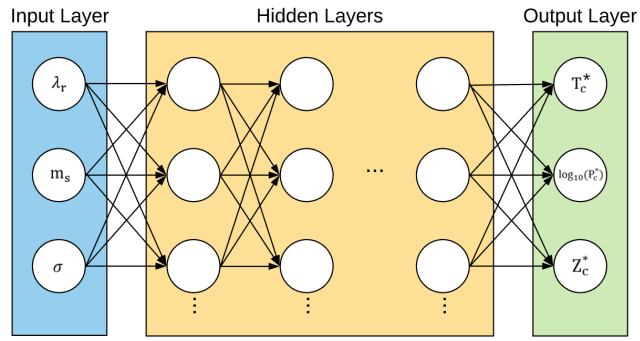


Fig. 1. Overall structure of an artificial neural network (ANN).

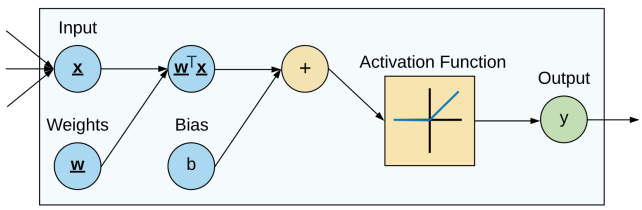


Fig. 2. Interior working of a single hidden unit.

In order to make predictions, an ANN must first be trained on suitable examples. The process of training is simply the optimisation of the weights at each node and the bias term of each layer to minimise an error function on known input/output examples of data. A back-propagation algorithm is then used to fit the weight and bias terms to the data. This algorithm is iterative and will adjust the parameters at each epoch by an amount proportional to the gradient of the error function with respect to each parameter.

III. Methodology

A. Data Collection and Pre-processing

As an input to this study, Yaw's Thermophysical Database for organic compounds was available [9]. This database contains experimental critical point data for more than 5,000 compounds, including the critical temperature, pressure and molar density. Using these parameters we can calculate the critical compressibility factor using:

$$Z_c = \frac{P_c}{RT_c\rho_c} \tag{4}$$

In addition to the experimental critical point data for each compound, the version of the database which we had access to also included estimations of the SAFT-VR Mie molecular descriptors for each compound which were obtained using a 4-parameter principle of corresponding states correlation. This uses the molecular weight, critical temperature, acentric factor and saturated liquid density (@ $T_r = 0.7$) as inputs to calculate molecular descriptors: the Mie segment length scale (σ), the Mie segment energy scale (ϵ) and the Mie potential repulsive constant (λ_r). It should be noted that according to London's theorem, it as assumed that all compounds have a Mie potential attractive constant of 6. The molecular descriptors for each compound were estimated multiple times for an assumed series of chain lengths with integer values from 1 to 6.

When training machine learning models, it is important to ensure that the data used for training is of the highest quality to prevent the final model learning any incorrect behaviour. To this end, we applied some pre-processing steps to Yaw's database to ensure that only the highest quality data was fed to our model for training. Firstly, we selected an appropriate value of the chain length for each molecule. This was necessary as only one of the assumed chain lengths per compound is actually representative of the true shape and size of the real molecule. To do this, the common heuristic was employed that the chain length should be approximately equal to the number of heavy (non-hydrogen) atoms in a molecule, divided by three. This selection process was performed using a Python script which interprets the chemical formula of a molecule to count the number of heavy atoms.

Next, we filtered out any compounds which had repulsive constants too small/large to be representative of physical behaviour. If the repulsive constant is less than 6, the repulsive forces in the potential are weaker than the attractive forces, meaning that the potential function would appear to be upside down - this is, of course, wrong. Also, values of the repulsive constant greater than approximately 35 are removed, as this is above the typical range of applicability for this molecular descriptor. Finally, we also removed any entries from the database which appeared to be erroneous. This included any negative/missing entries for any of the experimental critical properties or estimations of the molecular descriptors.

B. Model Training

As a preliminary investigation, we wanted to see if was even feasible for an ANN to produce meaningful predictions for critical properties. To do this, the ANN was initially trained using a large grid of pseudo-data generated using the SAFT-VR Mie EoS. This training data contained SAFT predictions of critical points for 20,000 "pseudo-compounds", the molecular descriptors for which were obtained by randomly sampling a log-normal distribution which was fitted to the distribution of molecular descriptors in Yaw's database. Once trained, both SAFT and the ANN were tested on a subset of 1,000 compounds from the Yaw database. The initial findings were promising, with the ANN almost exactly matching SAFT in terms of predictive accuracy, with performance metrics reported in Appendix A. This clearly shows an ANN is able to mimic the underlying theoretical description of critical points from SAFT. This approach however limits the performance of the ANN to that of SAFT, and perhaps the predictive accuracy could be improved if the model was trained solely on experimental data.

To this end, the entire pre-processed experimental Yaw database was split 80/20 into training and testing sets respectively. Note the testing set used here is the same as the testing set used in the preliminary investigation, to quantify the impact of the different training sets on predictive accuracy. The training set was passed to the ANN Adam Optimiser to allow the learning of the weights and bias terms. The features

used from the training set were the segment length scale (σ), the chain length (m_s) and the Mie potential repulsive constant (λ_r). Feature scaling and mean normalisation was applied to each of these features to ensure they all roughly lie in a range of [-1, 1] across the set of training examples, according to the formula:

$$\mathbf{f_{scaled}} = \frac{\mathbf{f_{raw}} - \mu}{sd} \tag{5}$$

Where **f** is a feature vector, μ is the mean of that feature in the training set and sd is the standard deviation in the training set. This ensures that all features are given equal importance in the arithmetic of the ANN as they are all on an equivalent order of magnitude.

These training features were fed as inputs to predict non-dimensional critical properties as outputs. The non-dimensional critical properties predicted were: critical temperature (T_c^*), log of critical pressure ($log_{10}(P_c^*)$) and critical compressibility factor (Z_c^*). These properties are related to the dimensional critical properties using:

$$T_c^* = \frac{T_c}{\epsilon} \tag{6}$$

$$P_c^* = \frac{P_c \sigma^3}{\epsilon} \tag{7}$$

$$Z_c^* = Z_c \tag{8}$$

C. Hyperparameter Optimisation

The optimal set of hyperparameters for the ANN were found via an exhaustive search over a specified grid of trial hyperparameters. For each distinct combination of hyperparameters, the modelling accuracy is calculated, characterised by the R^2 value, and the combination with the greatest R^2 is selected for final testing. Here, each grid search uses the entire training set from Yaw's database, with k-fold cross validation using $k = 5$. The hyperparameters which were optimised were the number of hidden layers, the number of hidden nodes per hidden layer, the form of the activation function and the value of the regularisation parameter, α, which is a parameter intended to prevent overfitting to the training data. The results of this optimisation procedure can be found in Appendix B. Both the hyperparameter optimisation, and implementation of the ANN were conducted in Python using the Scikit-Learn machine learning package [10].

D. Performance Metrics

Three performance metrics are used to evaluate the predictive performance of the ANN and SAFT for compounds in the testing set. The first is the R^2 value, which evaluates the variance between predictions and experimental values.

The root mean square error (RMSE) is also used, and is calculated for the critical temperature using:

$$RMSE = \sqrt{\frac{\sum_i^N (T^*_{c,\mathrm{pred,i}} - T^*_{c,\mathrm{exp,i}})^2}{N}} \tag{9}$$

where $T^*_{c,\mathrm{pred}}$ is the predicted critical temperature, $T^*_{c,\mathrm{exp}}$ is the experimental critical temperature and N is the number of samples in the testing set. Finally, the mean absolute percentage error (MAPE) is evaluated. Again, for the critical temperature this is:

$$MAPE = \frac{1}{N} \sum_i^N \frac{|T^*_{c,\mathrm{pred,i}} - T^*_{c,\mathrm{exp,i}}|}{T^*_{c,\mathrm{exp,i}}} \times 100\% \tag{10}$$

Analogous equations to (9) and (10) are used to evaluate RMSE and MAPE for the predictions of critical pressure and critical compressibility factor.

E. Model Testing

We present two testing cases in the results and discussion section: firstly, we test the performance of the ANN for 1,000 organic compounds not included in the training set. Secondly, we test the performance of the ANN on an alternative database of 244 inorganic compounds, to gain insight of the capability of the model to make accurate predictions on unfamiliar families of compounds. In both testing cases, we also use the SAFT-VR Mie EoS as a reference case against which to compare the performance of the ANN.

IV. Results and Discussion

A. Organic Testing Set

Tables I and II show the predictive performance metrics for the ANN and SAFT, respectively. When predicting critical temperature, both SAFT and the ANN show very high prediction accuracy, both achieving an R^2 value of 0.999 and MAPEs of 0.26% and 1.23%, respectively. This behaviour is to be expected from SAFT, as the molecular descriptors for each compound are fitted to the known critical temperature. The ANN achieving similar levels of accuracy suggests the neural network is able to effectively mimic the theoretical description of critical temperature from SAFT. Another benefit of the ANN is, once trained, any predictions made amount to simple matrix multiplication, making the prediction time orders of magnitude faster than the iterative optimisation approach used in SAFT. In our testing this amounted to at least a 100x reduction in prediction time, but this will vary depending on the specific computer hardware used. Plots for the experimental vs predicted critical temperature for each model can be found in Fig. 5 and Fig. 6 in Appendix C.

Overall, predictions for critical pressure are worse than those for critical temperature and show a larger separation in modelling accuracy between SAFT and the ANN. The ANN outperformed SAFT by a reasonable margin, reaching an R^2 of 0.908 and an MAPE of 7.56% compared to SAFT's R^2 of 0.600 and MAPE of 16.9%. Plots for the predicted vs experimental critical pressure are shown in Fig. 3 and Fig. 4, which include distributions for the predicted and experimental values, a blue line of best fit and a black dotted 45° line. The spread for the ANN is much smaller and captures the trend more closely, with the line of best fit lying closer to the 45° line. At higher pressures we also see some banding

TABLE I
ANN PREDICTIVE PERFORMANCE METRICS (ORGANICS)

	Critical Temp	Critical Pressure	Compressibility Factor
R^2	0.999	0.908	0.539
RMSE (-)	0.031	0.005	0.016
MAPE (%)	1.23	7.56	4.36

TABLE II
SAFT PREDICTIVE PERFORMANCE METRICS (ORGANICS)

	Critical Temp	Critical Pressure	Compressibility Factor
R^2	0.999	0.600	-7.62
RMSE (-)	0.020	0.011	0.069
MAPE (%)	0.26	16.9	24.2

in the predictions from SAFT, where it predicts the same critical pressure for a range of different compounds. From closer observation of these data points, many compounds have either identical or very similar molecular descriptors to one another and it seems SAFT is unable to account for these small differences when making its predictions. This behaviour is not shown by the ANN, which exhibits no banding, and may be attributed to the supervised training stage. When the ANN is trained, it is shown many examples of inputs and the associated outputs, allowing it to "learn" the effect of small differences in the molecular descriptors on the critical pressure.

Outliers from each model are identified in Fig. 3 and 4 as A, B and C, and were investigated further to try and understand the poor modelling performance. Outlier A corresponds to decamethyltetrasiloxane ($C_{10}H_{30}O_3Si_4$), a long chain organic with a calculated chain length of 6. Looking through the entire Yaw database, only 23 of the 5,000 compounds have a chain

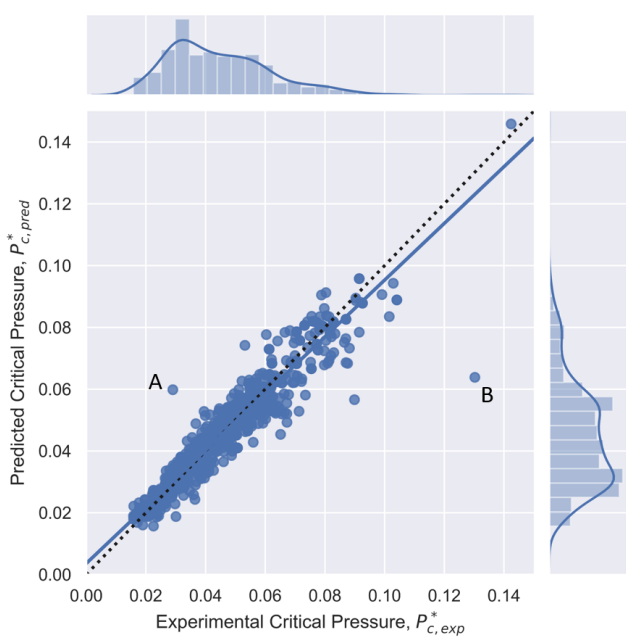


Fig. 3. ANN predictions vs experimental values pressure

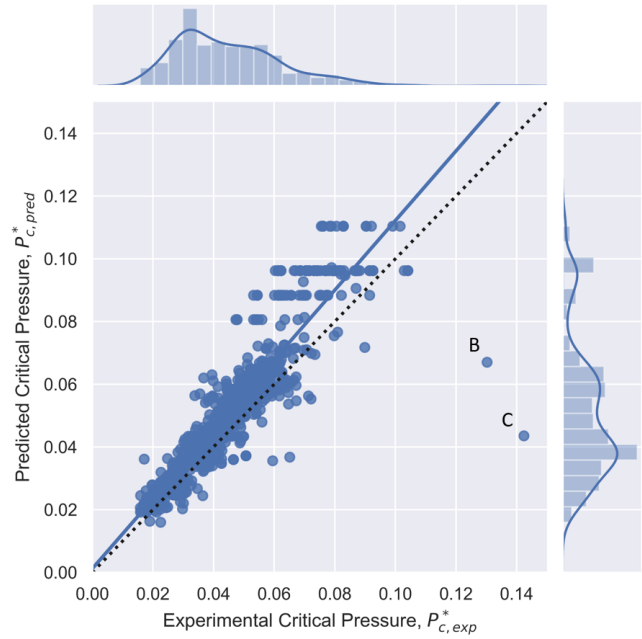


Fig. 4. SAFT-VR predictions vs experimental values pressure

length of 6. This suggests there aren't enough data points in the training set for the ANN to accurately predict critical properties of these long organics. This demonstrates a key limitation of machine learning methods; they do not generalise well if the testing points are not well represented within the training set.

Outlier B is present in both the ANN and SAFT predictions, with the predicted critical pressure from both models having an error in excess of 100%. This point corresponds to methylpropylsulphate ($C_4H_{10}SO_4$), a relatively short organic, which should be well represented by the training set. The large predictive errors could be related to its molecular structure, as it is an asymmetric compound, with a large electronegative center arising from the sulphate group. The molecular descriptors used in SAFT do not capture the polarity of molecules well, which may be why we see this poor predictive accuracy. Outlier C, dibutyl sulphate ($C_8H_{18}SO_4$), shows a similar pattern with a large prediction error of over 300% and the presence of central electronegative sulphate group. In future work this could potentially be overcome by including the dipole moment of each molecule as a training feature.

Predictions of compressibility factor show an even larger separation in modelling accuracy between the two approaches. SAFT produces predictions with essentially no correlation, reflected by an R^2 value of -7.62 and an MAPE of 24.2%. Conversely, the ANN performs much better, with an R^2 of 0.539 and an MAPE of only 4.36%. The difference in performance is clear from Fig. 7 and 8 in Appendix C, with most ANN predictions clustered around the 45° line with a relatively small spread. On the other hand, SAFT largely overestimates the compressibility factor for the majority of compounds, and carries forward the banded behaviour from the predictions of critical pressure. These results are very promising, showing that a relatively simple ANN can simultaneously predict all

three critical properties and do so to a much higher accuracy than the SAFT-VR Mie EoS.

B. Inorganic Testing Set

TABLE III
ANN PREDICTIVE PERFORMANCE METRICS (INORGANICS)

	Critical Temp	Critical Pressure	Compressibility Factor
R^2	0.957	-132	-0.44
RMSE (-)	0.177	0.313	0.060
MAPE (%)	8.06	39.14	14.2

The ANN performed much worse when predicting properties for unfamiliar inorganic compounds. It does produce reasonable predictions for critical temperature, but suffers greatly for predictions of both critical pressure and critical compressibility factor. Looking through the testing data used, this seems to be again due to a significant mismatch between the training and testing datasets. Many inorganics in the testing set have experimental critical properties massively outside of the ranges covered in the training set. Also, over 100 compounds have a chain length of 1, while only 26 compounds in the organic training set have a chain length of 1. Furthermore, many of the inorganics are either have permanent dipole moments or exhibit hydrogen bonding, both of which are not represented by the set of SAFT-VR Mie molecular descriptors used as training features. Future work could investigate splitting the entire inorganic dataset into a training and testing set to overcome this issue, as shown for the previous section for organic compounds. However, it should be noted that the available inorganic set contains only 244 compounds, and so modelling accuracy may decrease as a result of the smaller available training set.

V. CONCLUSION

In this work, we have presented an ANN able to predict the critical temperature, critical pressure and critical compressibility factor, using molecular descriptors of chain length, Mie repulsive constant and Mie segment length scale as training features. Predictions for organic compounds are presented, and compared with predictions from the SAFT-VR Mie equation of state. An ANN has shown multiple improvements over the SAFT-VR Mie EoS, summarised below:

- An ANN is able to simultaneously predict all three critical properties, and do so to a greater degree of accuracy compared to SAFT
- Large reduction in computational time for predictions compared to SAFT ($\sim$100 times)
- An ANN does not require an initial guess for the critical volume or range of permissible values for the critical volumes as the calculations are not iterative, which reduces user error
- The methodology shown is easily reproducible using open-source machine learning packages available in Python

Although these improvements are significant, there are some downsides to the proposed approach. The ANN provides only point predictions of the critical properties, and much like experimental data, predictions do not carry significant meaning without some indication of their error. This could be overcome through the use of other machine learning methods such as Gaussian Process Regression, which can provide probabilistic predictions and therefore give error bounds for predicted values.

The ANN also has poor predictive accuracy for compounds which are not represented well in the training set, which therefore limits its ability as a practical predictive tool for unknown compounds. The inherent circular nature of the method, needing knowledge of the critical temperature *a priori* to calculate the molecular descriptors, also presents problems for practical application. This could be addressed in further work by correlating more readily available molecular descriptors, such as molecular weight, dipole moment, boiling point and saturation pressure, with the set of critical properties. If successful, this would provide a truly predictive tool, able to make accurate critical property predictions for potentially any unknown compound.

REFERENCES

[1] N. et al., "Estimation of pure component properties part 2. estimation of critical property data by group contribution," *Fluid Phase Equilibria*, vol. 252, pp. 1–27, 2007.

[2] T. Lafitte, A. Apostolakou, C. Avendaño, A. Galindo, C. S. Adjiman, E. A. Müller, and G. Jackson, "Accurate statistical associating fluid theory for chain molecules formed from mie segments," *The Journal of Chemical Physics*, vol. 139, no. 15, p. 154504, 2013.

[3] T. K. Ulrich Deiters, *High-Pressure Fluid phase equilibria, Volume 2*. Elsevier, 2012.

[4] J. R. M. A. Ahmadi, T. Kashiwao and A. Bahadori, "Accurate prediction of properties of carbon dioxide for carbon capture and sequestration operations," *Petroleum Science and Technology*, vol. 34, no. 1, pp. 97–103, 2016.

[5] M. T. a. N. A. Arash Shadravan, "Intelligent tool to design fracturing, drilling, spacer and cement slurry fluids using machine learning algorithms," *Society of Petroleum Engineers*, 2015.

[6] G. et al, "Prediction of the boiling and critical points of polycyclic aromatic hydrocarbons via wang-landau simulations and machine learning," *Fluid Phase Equilibria*, vol. 484, pp. 225–231, 2019.

[7] A. Ervik, A. Mejia, and E. A. Müller, "Bottled saft: A web app providing SAFT-γ mie force field parameters for thousands of molecular fluids," *Journal of Chemical Information and Modelling*, vol. 56, no. 9, pp. 1609–1614, 2016.

[8] K. Zhu. (2019) SAFT-VR Mie open-source python implementation. https://github.com/kezhengzhu/scriptables/tree/master/saft.

[9] C. L. Yaw. (2010) Yaws thermophysical properties of chemicals and hydrocarbons. https://app.knovel.com/web/toc.v/cid:kpYTPCHE02/viewerType:toc/.

[10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

Electro-oxidation of Glycerol with Pt and Pd Deposition

Kathryn Jaitly, Athar Kamran
Department of Chemical Engineering, Imperial College London, U.K.

Abstract

Glycerol electro-oxidation has been shown to reduce the required level of input energy for the electrocatalytic production of hydrogen fuel by over 50%. However, to make the process economically viable, the reaction mechanism must be fully understood to optimise selectivities towards high value products. In this study, we effectively report the performance towards glycerol electro-oxidation reaction (GEOR) of platinum (Pt) and palladium (Pd) electrodeposited on glassy carbon substrates. A range of surface characterisation techniques are used to identify optimum surface morphology towards GEOR. The influence of sulfuric acid concentration was also investigated, and theories postulated for a proton-driven mechanism. It can be seen that dropcasted ink outperforms electrodeposition. However, the formation of a Pt film rather than nanoparticle formation hindered the ability of glycerol to adsorb onto the electrodeposited surface. Pt ink was used to carry out chronoamperometry analysis in order to identify conditions for achieving a 100% selectivity towards a valuable derivative: glyceraldehyde. Furthermore, the parameters of voltage, initial glycerol concentration and reaction time also play a crucial role in the reaction pathway.

1 Introduction

Energy demand for the transportation industry is estimated to rise by more than 25% by 2040 as a direct result of improved economic activity, increase in commercial transportation and rising aviation demand (ExxonMobil, 2019). To meet these needs, the transportation industry relies on carbon intensive fossil fuels. 22% of global CO_2 emissions are attributed to the transport sector (Rodrigue et al., 2016). Therefore, clean and efficient energy solutions must be investigated to meet these energy challenges and limit CO_2 emissions.

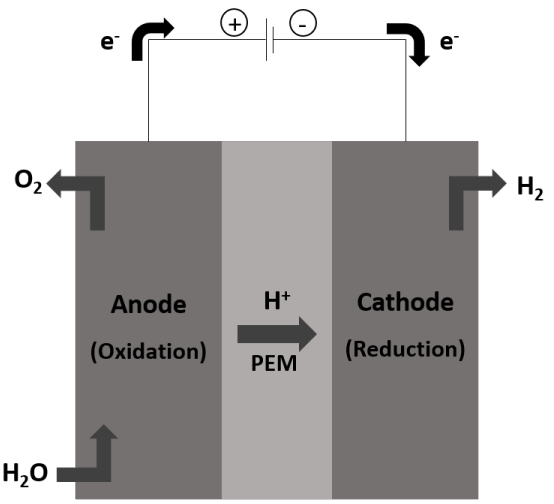


Figure 1.1: Electro-oxidation of water for H_2 generation

Hydrogen fuel produced via electrolysis of water has gained a strong degree of interest. As described in the electrolyser system in figure 1.1, H_2O enters the system and is oxidised at the anode resulting in the release of H^+. These protons are transferred across a proton exchange membrane (PEM) towards the cathode for reduction and H_2 formation. However, the typical input potential for the process is $E_{H_2O,EC}$ = 1.9V, as depicted in figure 1.2 (Simões et al., 2012). This leads to electricity consumption of approximately 4.8 kWh.$(Nm^3H_2)^{-1}$. The process has a poor cost efficiency thus restricting global applicability.

As reported in Marshall and Haverkamp (2008), alcohols are oxidised at lower potentials than water, therefore consume less energy. This is depicted in figure 1.2 where the use of glycerol instead of water as a reactant decreases the cell voltage for hydrogen production to $E_{gly,EC}$ = 0.7V. This decrease reduces consumption by over 50%, lowering the process input energy to 2.4 kWh.$(Nm^3H_2)^{-1}$. This aspect makes the glycerol electro-oxidation reaction (GEOR) a very attractive process.

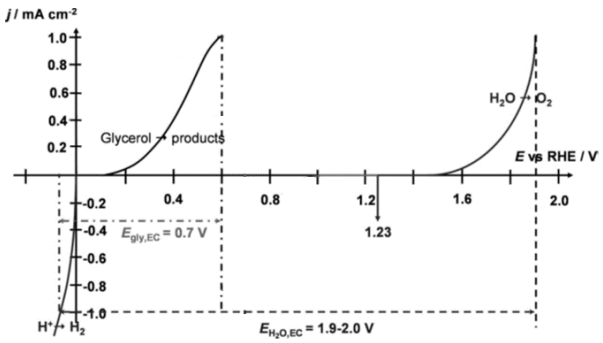


Figure 1.2: Butler–Volmer kinetics law for water oxidation and glycerol oxidation

However, to account for the economic viability of the process, it is important to think about the markets surrounding glycerol, its derivatives and H_2 fuel produced. Glycerol is produced as a waste by-product in large quantities during biodiesel formation and soap manufacture. This excess accumulation has allowed glycerol price to fall below 0.25p/g (Sigma Aldrich). A level 1 cost analysis, focused on the current prices of the reactant utilised and product formed, can be carried out to assess the economic potential of the process at this initial stage Douglas (1988).

The complete oxidation of glycerol to CO_2 is presented as follows and allows for the maximum theoretical H_2 production:

$$C_3H_8O_3 \ + \ 3H_2O \ \rightarrow \ 3CO_2 \ + \ 7H_2$$

Alternatively, the minimum theoretical H_2 production follows the glycerol partial oxidation mechanism with glycerol conversion to glyceraldehyde:

$$C_3H_8O_3 \ + \ H_2O \ \rightarrow \ C_3H_6O_3 \ + \ H_2$$

Assuming a 100% conversion of glycerol, complete oxidation produced 0.28g of H_2/g of glycerol reacted, while partial oxidation produced only 0.04g of H_2/g of glycerol reacted. However, when the economic potential was calculated and compared using price metrics as stated in the supporting information (SI) (section A.1), it was found the profits from the partial oxidation mechanism were higher than with complete oxidation. This can be attributed to the wide array of applications for glyceraldehyde in cosmetic and medical industries which is reflected in its market share value of £30.60/g (Sigma Aldrich).

GEOR has been studied in great depth and literature spans back over 25 years (Avramov-Ivic et al., 1993). The mechanism for the oxidation reaction has been accepted to follow the pattern as shown in figure 1.3. However, the combination of glycerol oxidation and hydrogen co-generation is a novel concept which has been developed in the past five years. A seminal paper in this area was Kim et al. (2014), who developed a 'glycerolyser' with PEM technology, allowing continuous glycerol throughput and hydrogen production. To optimise the process further and to increase profitability, research has to focus on the glycerol partial oxidation mechanism in order to optimise for economic feasibility. There are several aspects of the cell that can be changed, including electrode material (Arjona et al., 2017), voltage, initial glycerol concentration (Kim et al., 2014), and pH(Kwon et al., 2011). However, to the best of knowledge, two routes have not been explored: the electrodeposition of metals on glassy carbon (GC) plates, and the effect of electrolyte concentration on cell operation. Papers have come close to replicating these tests: in Arjona et al. (2017), the influence of Pt and Pd electrodeposition on carbon paper and nanofoam was examined. Meanwhile, Kwon et al. (2011) compared platinum and palladium in acidic, alkaline and neutral conditions - but did not compare different acidities; similarly, Zhang et al. (2012) considered systems at varying alkalinities - but did not consider the effect of acidity.

Based on this literature review, the paper was condensed to two areas of research: the electrodeposition of platinum and palladium, and the effect of electrolyte concentration on the extent of reaction. Platinum and palladium were chosen on the basis of Arjona et al. (2017). Platinum exhibits higher catalytic activity in acidic media, but palladium showed higher tolerance to CO catalytic poisoning. Electrolyte concentrations were chosen based on Zhang et al. (2012). Sulfuric acid was chosen due to its use throughout literature in place of $HClO_4$ (Kim et al., 2014), (Kwon et al., 2011), (Houache et al., 2019). Several papers have carried out GEOR in 0.5M H_2SO_4, but no literature has tried experimentation in stronger concentrations of sulfuric acid (Houache et al., 2019).

Based on these initial choices, the project was planned as follows. Platinum and palladium were deposited via two methods: electrodeposition, and dropcasting. Ink dropcasting is typically used as a control to compare results (Houache et al., 2019). The surfaces were characterised using microscopy, XRD and XPS. Cyclic voltammetry (CV) was then carried out to measure the activity of each deposition; the most active application method was carried forward to the second round of testing: chronoamperometry (CA). CA was carried out for eight hours at different voltages, glycerol concentrations and sulfuric acid concentrations. Samples were taken every two hours to monitor the progress of the reaction, and were analysed using HPLC to measure product concentration.

This study reports the subsequent results of these variable parameters and discusses the pos-

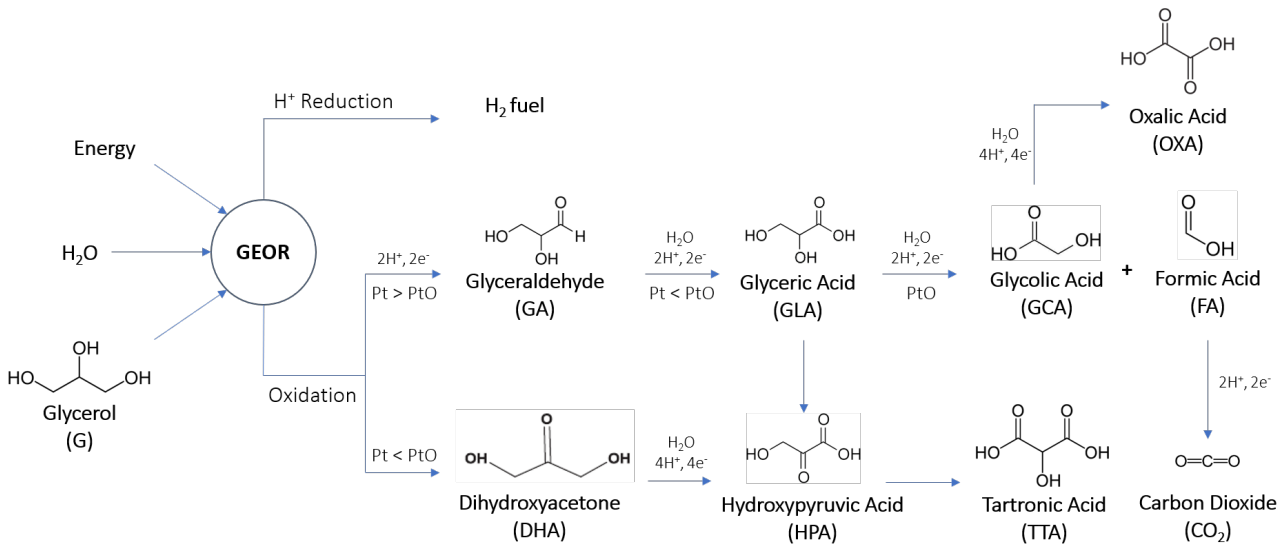


Figure 1.3: Proposed reaction pathways for glycerol electro-oxidation, adapted from (Kwon et al., 2011)

sible influences on the reaction pathway.

2 Methodology

2.1 Deposition

2.1.1 Electrodeposition

Platinum electrodeposition was carried out based on Duarte et al. (2006) in a three-electrode system. 25mL of 2mM hexachloroplatinic acid solution in 2M H_2SO_4 was used as an electrolyte. Two consecutive potential steps were applied and repeated for ten cycles using a Multi Autolab Potentiostat (M101). The first step was applied at -0.469V (vs. Ag/AgCl) for 5s followed by a step at 0V (vs. Ag/AgCl) for 60s.

Palladium electrodeposition was carried out using a template bubble deposition method, performed on a two-electrode system Winiwarter et al. (2019). 25mL of 2mM palladium chloride solution in 2M H_2SO_4 was used as an electrolyte and Pt foil was used as a counter electrode. Voltage was applied at -1V (vs. counter electrode) for 60s, then at -4V (vs. counter electrode) for 180s using a Biologic Potentiostat (VSP-300). Both electrodes were thoroughly rinsed with DI water and allowed to dry before further use.

2.1.2 Ink dropcasting

Platinum catalyst ink was prepared based on an internal Titirici Group method. 4mg of commercial 20wt/% Pt/C powder (Sigma Aldrich) was mixed with 700μl of deionised water. 264μl of ethanol was added dropwise and the mixture sonicated for 30s. 36μl of Nafion (D-521 dispersion 5% w/w in water and 1-propanol supplied by Alfa Aesar) was added as a binder and the mixture was re-sonicated in a bath for 10 minutes. For effective catalyst distribution, probe ultrasonication was carried out on the catalyst/C-Nafion ink using a Fisherbrand^{TM} Q700 Sonicator for 10 minutes at an amplitude of 40μm. xμl (where x = 300 or 600) of the ink was drop-casted onto a freshly polished 25 x 25mm^2 GC plate and allowed to dry in a fume hood overnight.

Palladium catalyst ink was prepared based on Verma et al. (2019). 20mg of commercial 20wt/% Pd/C catalyst (Sigma Aldrich) was mixed with 800μl of deionised water. 800μl of isopropyl alcohol was added dropwise and the mixture sonicated for 30 seconds. 52μl of Nafion (D-521 dispersion 5% w/w in water and 1-propanol supplied by Alfa Aesar) was added as a binder and the mixture was re-sonicated in a bath for 10 minutes. For effective catalyst distribution, probe ultrasonication was carried out on the catalyst/C-Nafion ink using Fisherbrand^{TM} Q700 Sonicator for 10 minutes at an amplitude of 40μm. yμl (where y = 200 or 400) of the ink was dropcasted onto a freshly polished 25 x 25mm^2 GC plate and allowed to dry in a fume hood overnight.

2.1.3 Cleaning

Electrodeposited plates were cleaned in aqua regia (1:3 volumetric ratio of fuming sulfuric acid and concentrated nitric acid) as described in Tegou et al. (2010). The resting time was increased to ten minutes due to the larger area

of the plate. The plate was then thoroughly washed in DI water and cleaned with acetone.

Ink-coated plates were initially cleaned using deionised water and acetone, then polished with Buehler MasterPrep^{TM} 0.05 μm Alumina on a polishing cloth. Plates were polished for one minute using a figure-of-eight motion. The surface was wiped down using water and acetone.

2.2 Catalyst Characterisation

2.2.1 Microscopy

SEM: GC plates were desiccated and attached to stubs using conductive tape or silver glue (Sigma Aldrich). Field-emission gun scanning electron microscopy (FEG-SEM) was performed using Leo Gemini 1525 (Zeiss, Germany) at 5kV.

TEM: Samples were prepared by dispersing deposition scrapings in ethanol and ultrasonicating for 5 minutes. Samples were dropcasted onto lacy carbon grids (Agar Scientific) and dried overnight in a desiccator. A FEI TITAN 80/300 TEM instrument at an accelerating voltage of 300kV was used to obtain high resolution transmission electron microscopy (HRTEM) images in TEM mode.

2.2.2 XRD

XRD was carried out using a PANalytical X'Pert Pro Multi Purpose Diffractometer using Cu Kα radiation (Kα = 1.5186Å). Patterns were recorded for 2θ between 10.0° and 90.0°.

2.2.3 XPS

X-ray photoelectron spectroscopy (XPS) analysis of the samples was performed with ThermalFisher Nexsa using an Al Kα1 x-ray source.

2.3 Electrochemical Experiments

2.3.1 Cyclic Voltammetry

Electroanalysis of the modified GC plate was carried out using cyclic voltammetry (CV) analysis. Experiments were performed on a typical three-electrode batch system. The modified GC plate acted as the working electrode, Pt foil was used as a counter electrode and an Ag/AgCl in 3M KCl solution was used as a reference electrode. The GC plate was connected to a stub using silver paint to ensure high electrical conductivity and wrapped in Teflon tape to prevent direct contact of the electrical circuit with the electrolyte.

20mL of 1M H_2SO_4 solution was syringed into the cell. Nitrogen was bubbled into the solution for 15 minutes. CV was carried out at a scan rate of 50mV/s in the voltage window [0V, 1.4V] (vs. RHE) for 20 cycles. The system was emptied and refilled with 20mL of 1M glycerol in 1M H_2SO_4 and the process repeated.

2.3.2 CA

Constant-voltage chronoamperometry (CA) experiments were carried out in the same three-cell batch set-up described earlier. Electrolyte concentration was varied as xM H2SO4 + yM glycerol where x = 1M, 2.5M and y = 0.2M, 1M. The solution was bubbled with N_2 for 15 minutes initially. Constant potentials of 0.75V, 0.90V, 1.05V and 1.20V (vs. RHE) were applied for 8 hours and N_2 was bubbled over the course of the experiment. 500μl samples were collected every two hours for product analysis.

2.4 Product Analysis

The reaction products were analysed by using an Aligent Technologies 1260 Infinity II HPLC equipped with an Aminex HPX-87H column, specific to alcohols, ketones, aldehydes and organic acid analysis. Chromatography was performed with a 0.01M H_2SO_4 aqueous solution as the mobile phase at a 0.6mL/min flow rate. The UV detector was set at λ=254nm and temperature of 65°C. The calibration was performed by using dl-glyceraldehyde (+90%), dihydroxyacetone (97%), tartronic acid (+97%), oxalic acid (+99%), glycolic acid (+99%), glyoxylic acid (98%) and glycerol (+99.5%) purchased from Sigma–Aldrich.

3 Results and Discussion

3.1 Material Selection

3.1.1 XPS

XPS was carried out successfully for three of four deposition techniques (see figure 3.1). Each method shows a platinum ($Pt4f_{5/2}$, $Pt4f_{7/2}$) or palladium peak ($Pd3d_{3/2}$, $Pd3d_{5/2}$), indicating that the correct metal was deposited. Signal deconvolution shows that each sample presents three different Pt species: Pt(0) at 71.6eV, Pt(II) at 72.6eV and Pt(IV) at 76eV. Based on

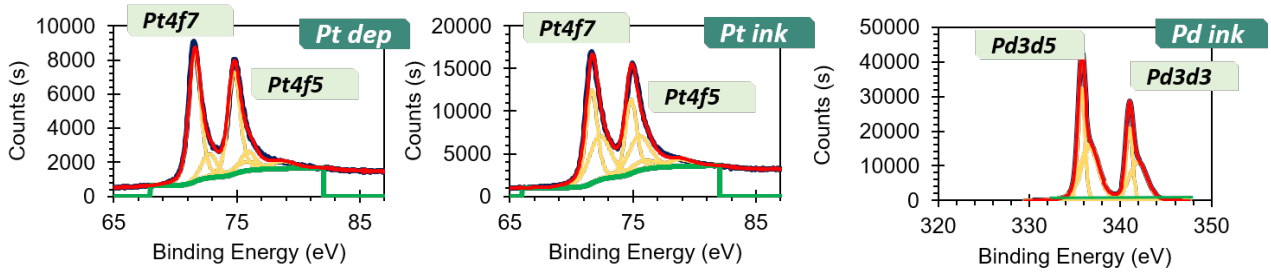


Figure 3.1: X-ray photoelectron spectroscopy (XPS) Analysis. Pd ink with corrected intensity counts

peak shape, Pt is dominated by Pt(O) rather than ions - agreeing with results presented in later sections. Two Pd species are present in the powder sample: Pd(0) at 335.4eV and Pd(II) at 336.8eV.

3.1.2 XRD

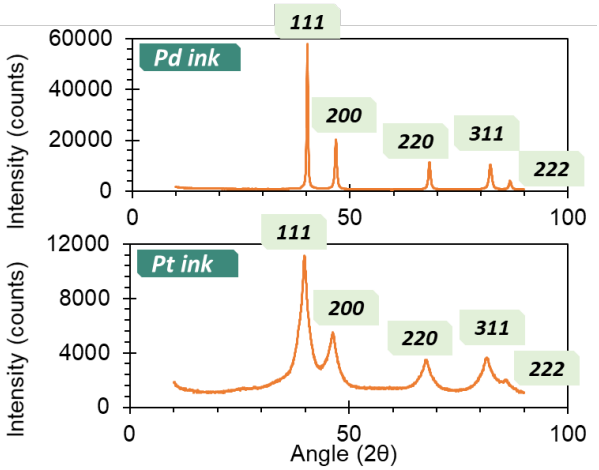


Figure 3.2: X-Ray Diffraction (XRD) Analysis

Platinum and palladium ink powders were examined via XRD. Results are plotted in figure 3.2. Peaks match with literature. However, platinum exhibits broad peaks whereas palladium exhibits sharp peaks. The reason for this is due to the nature of the powders. Platinum is sputtered onto an amorphous carbon support, causing a large range of particle size to form, causing a broader peak. The palladium powder has a crystalline structure with consistent particle size, therefore sharper peaks.

3.1.3 Microscopy

Figure 3.3 shows the appearance of Pt electrodeposition using SEM and TEM. SEM microscopy identified that the majority of the surface was covered by a film of platinum, rather than small nanoparticles seen in the work of Duarte et al. (2006). The lack of nanoparticles is further emphasised through the TEM image, which shows a bulk polycrystalline metal. Whilst the method was adapted to include multiple pulses, this should not cause the presence of a thin film on the surface, as this methodology should control particle size rather than create a film. It can also be seen that many papers cite this paper, indicating the reliability of the method (Arroyo-Gómez et al., n.d.), (Afsham et al., 2019), (Fischer et al., 2018).

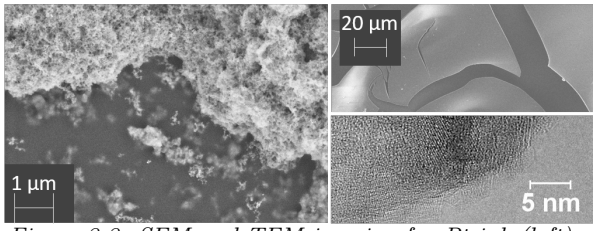


Figure 3.3: SEM and TEM imaging for Pt ink (left) and Pt electrodeposition (right)

It is therefore suggested that the method was incorrectly adapted to our set-up. Rather than convert the second pulse voltage to 0V vs. RHE, it was decided to maintain a voltage of 0V vs. Ag/AgCl. The result of this is a negative current is obtained in the second pulse of the voltammetry. As can be seen in Rajalakshmi and Dhathathreyan (2008) the pulse should have no current. However, it is difficult to judge if this is the reason behind film formation without carrying out further experiments.

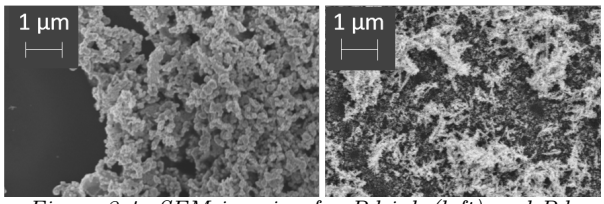


Figure 3.4: SEM imaging for Pd ink (left) and Pd electrodeposition (right)

Palladium electrodeposition shows expected morphology, as the SEM image (figure 3.4) resembles that of Winiwarter et al. (2019).

3.1.4 CV

Cyclic voltammetry is a test over a range of voltages to find regions of high current, indicating high catalytic activity. It is characterised by two scans - anodic (increase in voltage to a maxi-

mum) and cathodic (returning to an initial voltage). In this display method, positive peaks will represent oxidation and negative peaks will represent reduction (Elgrishi et al., 2017). A typical glycerol reaction curve consists of three peaks: one oxidation peak during the anodic scan, and two oxidation peaks during the cathodic scan. This represents the stages of the platinum surface. At low voltages during the anodic scan, glycerol adsorbs onto the surface. This limits further glycerol adsorption, leading to a higher onset potential. Current density is high due to high rate of reaction. During the cathodic scan, platinum oxides are reduced and the surface is cleaned, allowing for further glycerol oxidation. This is a new concept in electrocatalysis, and the entire concept is outlined in Zhao et al. (2016).

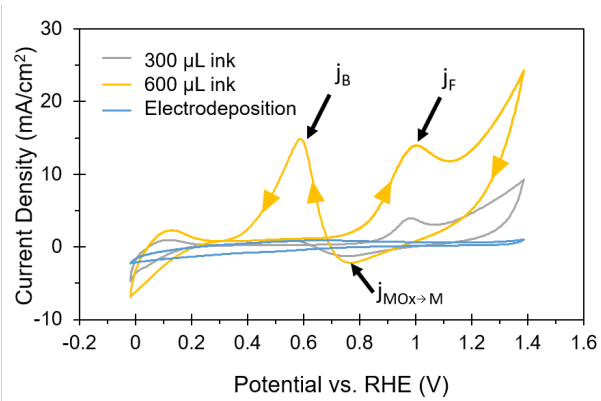


Figure 3.5: Cyclic Voltammetry analysis on Pt

In figure 3.5, platinum ink shows the highest catalytic activity due to the presence of the largest peak – 600μL. 300μL is also active, with three distinct peaks visible. 600μL is more active in this case because a larger mass of platinum is deposited on the surface, therefore allowing a larger amount of catalytic activity. Platinum electrodeposition exhibits no activity. As mentioned in section 3.1.3, electrodeposition forms a continuous film rather than nanoparticles. There is less ability for the glycerol to access the platinum surface. Nanoparticles are preferable due to their small size and large surface area, therefore exhibit larger electrocatalytic activity.

Similar cyclic voltammograms are seen in literature (Lee et al., 2019). It is important to note that peaks may differ based on the concentration ratio of materials, meaning that the reduction peak may look disproportionately small (Compton and Banks, 2011).

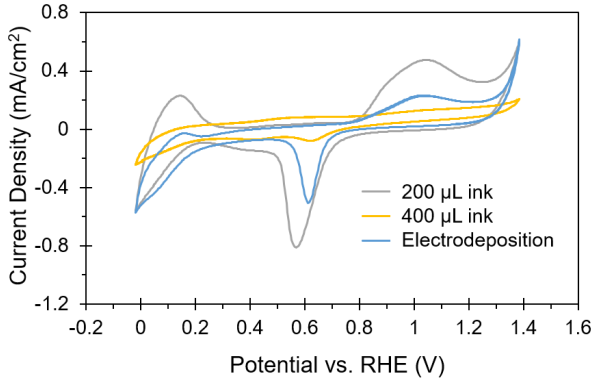


Figure 3.6: Cyclic Voltammetry analysis on Pd

Figure 3.6 shows voltammograms for all palladium materials. When comparing palladium and platinum cyclic voltammetry graphs, it can be seen that current density magnitude is low for all Pd materials. Palladium is a lot less active than platinum in acidic conditions. This is supported by literature, where palladium exhibits little activity in acidic media (Arjona et al., 2017), (Houache et al., 2019).

Pd voltammograms only show two peaks, representing a anodic glycerol oxidation peak and a cathodic reduction peak for metals. As it has been shown that palladium is inactive in acidic media, it is put forward that there is no second oxidation peak due to the lack of activity on palladium surface; at voltages which would potentially show glycerol activity on cleaned surfaces, reduction dominates therefore no additional glycerol peak is seen. It is noted that palladium was initially chosen due to its higher tolerance to catalytic poisoning by CO (Arjona et al., 2017), but this is not the case based on the work of (Zhao et al., 2016).

400 μL of ink is shown to cause a lower anodic peak than 200 μL. This could be due to such high masses of palladium on the surface that glycerol has limited access to the metal. Electrodeposition does show electrocatalytic activity. Whilst Pd electrodeposition could not be proven via XPS, the similarity of the peaks when compared to a commercial ink indicates the presence of Pd on the plate surface. Similar peaks are seen in literature (Zalineeva et al., 2015).

It is noted that to better characterise the effectiveness of each deposition method based on access, electrocatalytic surface areas should be calculated for the material. Methods for calculation are outlined in Wei et al. (2019).

Based on these CV results, it was decided to use Pt ink at 600 μL for section 3.2.

3.2 Chronoamperometry

3.2.1 The Impact of Voltage - 1M glycerol, 1M H_2SO_4

In figure 3.7, all lines start at a high current and decrease with time. This is due to electron transfer across the double layer. Initially, there is high current flowing across the double layer as electrons are driven to the cathode. leading to electrochemical reactions. Eventually, the number of glycerol molecules at the electrode surface decreases with time and transfer of molecules to the electrode surface becomes diffusion-limiting. A steady state is reached, causing the rate of reaction and current to be constant (Bockris and Amulya, 1970).

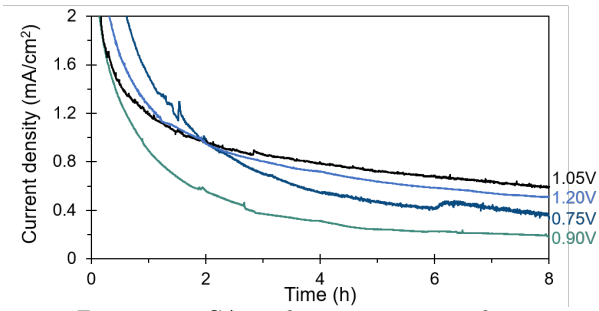


Figure 3.7: CA analysis at varying voltages

It is noted that there are fluctuations in the CA data. This noise is due to turbulence caused by nitrogen flow. Nitrogen bubbling was poorly controlled through the experiment due to excessive regulator size and flow control was manual, thus causing spurts of gas. This caused the solution to mix, allowing more glycerol to reach the surface due to advection. This allowed a larger rate of reaction, larger electron flow and therefore spikes in current.

When comparing gradients of CAs, it can be seen that lower voltages (0.74V and 0.89V) have 'deeper' curves, whereas higher voltages have shallower curves; at 8h, higher voltages finish at higher current densities. This could be due to platinum oxidation (Pourbaix et al., 1959) or surface cleaning at low potentials (Zhao et al., 2016) - but this is conjecture. However, it should be noted that there are some experimental limitations. Each plate was partially covered with parafilm, leading to estimates of total surface area exposed to solution and therefore of current density. Furthermore, due to limited time availability, each experiment was carried out once. From Winiwarter et al. (2019), chronoamperograms show variability - especially at voltages close to the peak potential. It is therefore important to note that these tests may not be repeatable.

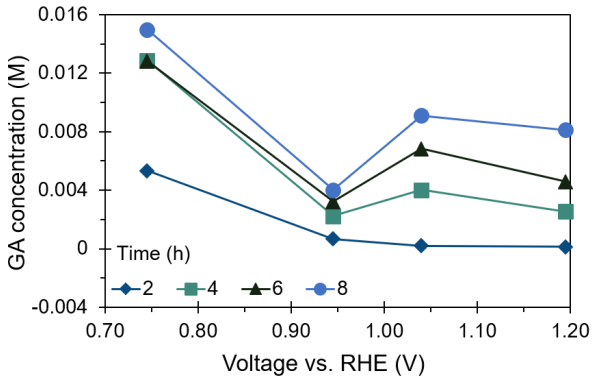


Figure 3.8: GA produced at 1M G and 1M H_2SO_4

Figure 3.8 shows the dependence of glyceraldehyde (GA) concentration on voltage, shown at different times throughout the experiment. The metric of GA was used because there is 100% GA selectivity. However, it is noted that other products may be present in small concentrations - calibration curves are included in the ESI (Section D) for further detail.

As time increases, GA concentration increases. Glycerol is converted to GA but cannot be oxidised any further. This is expected from the mechanism detailed by Kwon et al. (2011) and outlined in figure 1.3 - XPS tests from 0h and 8h indicate that only Pt is present on the plate, indicating that the only possible product route is towards GA.

GA concentration varies with voltage. The lowest GA concentration is seen at 0.89V, with larger GA yields at both lower and higher voltages. This dip coincides with the platinum reduction peak seen in figure 3.5. This reaction may dominate glycerol reaction. At higher voltages, GA yield increases. This is thought to be due to high voltages oxidising adsorbed CO on the Pt surface, cleaning the surface and increasing availability. Furthermore, these patterns are similar to trends presented by Kwon et al. (2011), which show high relative yields of GA in comparison to other products. However, it is noted that Kwon produces a range of other products. This could be for several reasons, including differing concentrations of electrolyte/glycerol, different sampling techniques, different cell geometries (altering length of diffusion pathway) and different plate areas (limiting current accessible to solution).

3.2.2 The Impact of Concentration

Figure 3.9 shows voltammograms for different concentrations of glycerol and sulfuric acid. It can be seen that all peaks change position slightly. This is particularly evident in 2.5M j_f peak, which is broader and has a smaller gradient at lower potentials. A lower onset potential implies that there could be more adsorbents at the surface at lower voltages, blocking access to the surface, or that the presence of H^+ blocks access to the plate, requiring a larger potential to overcome this charge.

The 0.2M G voltammogram has smaller oxidation peaks in comparison to 1M G voltammogram. This is due to smaller concentrations of glycerol, therefore less glycerol is at the surface and lower rates of reaction.

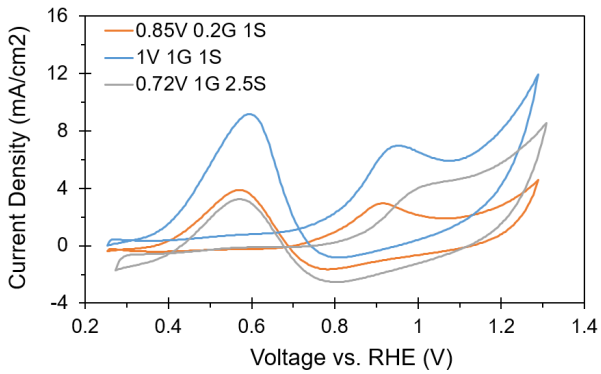


Figure 3.9: CV analysis at varying concentrations

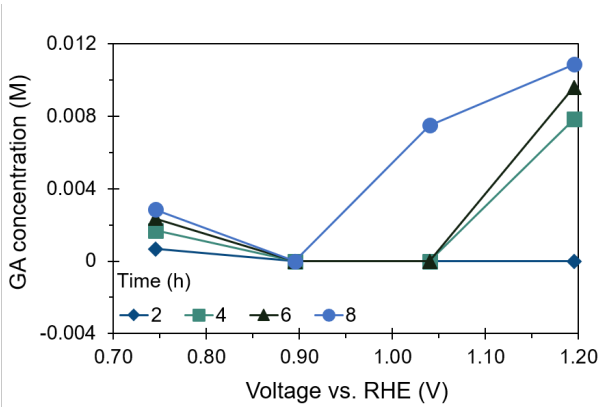


Figure 3.10: GA produced at 0.2M G and 1M H_2SO_4

The impact of glycerol concentration can be seen in more detail through CAs and HPLC analysis. Figure 3.10 shows the trend between glycerol oxidation and voltage. The overall pattern has not changed. Lowest GA concentration is seen at 0.9V, highest yield is seen at 8h and there is still 100% selectivity towards GA. Instead, we see changes in relative yields of concentration with voltage. At 0.75V, there is a low yield of GA. At concentrations above 0.9V, higher yields are produced, with similar concentrations to those obtained in 1M G, 1M H_2SO_4. The relative concentrations at higher and low voltage differ drastically between the two runs. It is suggested that voltages below 0.9V are glycerol-limited whereas voltages above 0.9V are limited by H^+ concentration.

It is noted that there is a contradiction with Kim et al. (2014), which states a change in selectivity with lower glycerol concentrations - less total glycerol results in less renewal at the surface and further oxidation beyond GA. This should be seen these results as Kim examines similar glycerol-to-platinum ratios to those discussed in this paper. The reasons for differences in selectivity are detailed in section 3.2.1.

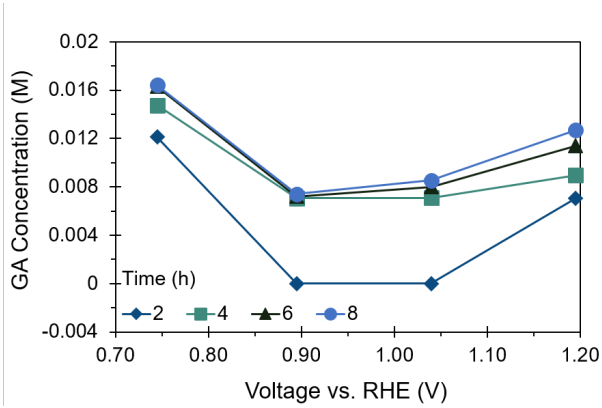


Figure 3.11: GA produced at 1M G and 2.5M H_2SO_4

The effect of electrolyte concentration can be seen in figure 3.11. We see a dip at 0.9V, and largest peaks at 0.75V and 1.2V - similar to section 3.2.1. However, the yields of GA at all voltages are higher than in other conditions. Glycerol conversion is also higher, as glycolic acid (GCA) is produced (see figure 3.12); GCA is produced at three of four voltages in this experiment, and concentration varies by voltage.

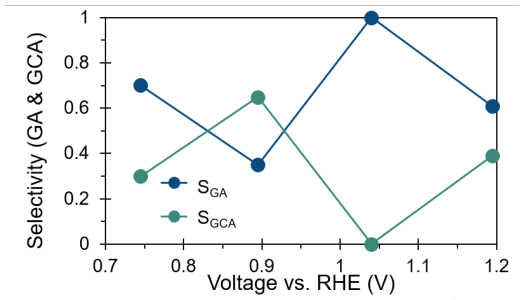


Figure 3.12: Selectivity towards GA and GCA at 1M G and 1M H_2SO_4

The increase in GA yield and GCA yield is thought to be driven by the presence of H^+, making products more stable. It may also be possible that the high concentration of H^+ acts as a charged barrier to the bulk solution, causing glycerol molecules to remain close to the plate

surface and react beyond GA. Glycerol oxidation to GCA indicates the presence of PtOx on the surface (see figure 1.3). XPS should be run on the surface to prove that PtOx is present in future experimentation. Furthermore, the presence of glycolic acid is unexpected. Based on Kwon et al. (2011), it is seen that GCA production is strongly base-catalysed, yet an increase in sulfuric acid concentration led to larger rates of GCA production in our experiments. It is suggested that sulfate ions are acting as bases during this experiment. However, it is difficult to judge based on the scope of the project.

Two limitations should be mentioned at this stage. Formic acid and hydroxypyruvic acid were not calibrated in HPLC. There were a couple of unidentified peaks in the spectrum whose area could not be calculated by the machine. It is possible that one of the unidentified peaks could represent the presence of either of these substances. The second caveat is regarding the surface itself. At such high concentrations, platinum dissolution is possible (Pourbaix et al., 1959). This set of conditions would therefore not be suitable for large-scale set-up, where stability is crucial. In future experiments, Pt ICP-MS should be used to calculate the Pt concentration in solution during the reaction, as well as the impact of the presence of PtOx (Kodera et al., 2007).

4 Conclusions

There are two areas of focus in this paper: metal electrodeposition and effect of sulfuric acid concentration. Generally, it can be seen that ink outperforms electrodeposition. However, the formation of a Pt film rather than nanoparticles hindered the ability of glycerol to react. Platinum ink was used to carry out chronoamperometry. Changes in voltage, concentrations and time are crucial to GEOR. 0.9V is shown to be relatively inactive to the reaction, increases in time aid conversion and the presence of stronger acids enable increased selectivity. 100% selectivity to GA was seen in 75% of the CA experiments, indicating that it is possible to achieve high yields.

5 Outlook

Eventually, the aim of the project is to create a functioning, optimised 'glycerolyser'. However, for optimum H_2 formation and reaction kinetics additional characterisations need to be carried out on these materials, such as ECSA calculations and XPS on all final samples. This should be incorporated into current and future studies. Platinum electrodeposition needs to be explored further to reduce metal use for future incorporation. Eventually, the aim of the project is to create a functioning, optimised 'glycerolyser'. The project will scale up to a H-cell to optimise parameters further.

References

Afsham, N., Fallah, N., Nassernejad, B., Javanbakht, M. and Jabbari, Z. (2019). Fabrication of gas diffusion electrode via pt electrodeposition on cathodic oxidized carbon paper as the anode for high-temperature polymer membrane fuel cell in the presence of co, *Ionics* pp. 1–12.

Arjona, N., Rivas, S., Álvarez-Contreras, L., Guerra-Balcázar, M., Ledesma-García, J., Kjeang, E. and Arriaga, L. (2017). Glycerol electro-oxidation in alkaline media using pt and pd catalysts electrodeposited on three-dimensional porous carbon electrodes, *New Journal of Chemistry* **41**(4): 1854–1863.

Arroyo-Gómez, J. J., Barrera, D., Castagna, R. M., Sieben, J. M., Alvarez, A. E., Duarte, M. M. E. and Sapag, K. (n.d.). Ptpdni catalysts supported on porous carbon for ethanol electro-oxidation, *ChemCatChem* .

Avramov-Ivic, M., Léger, J.-M., Beden, B., Hahn, F. and Lamy, C. (1993). Adsorption of glycerol on platinum in alkaline medium: effect of the electrode structure, *Journal of Electroanalytical Chemistry* **351**(1-2): 285–297.

Bockris, J. O. and Amulya, K. R. (1970). *Modern Electrochemistry: Volume 2. Fundamentals of Electrodics*, Springer.

Compton, R. G. and Banks, C. E. (2011). *Understanding voltammetry*, World Scientific.

Douglas, J. M. (1988). *Conceptual design of chemical processes*, Vol. 1110, McGraw-Hill.

Duarte, M., Pilla, A., Sieben, J. and Mayer, C. (2006). Platinum particles electrodeposition on carbon substrates, *Electrochemistry communications* **8**(1): 159–164.

Elgrishi, N., Rountree, K. J., McCarthy, B. D., Rountree, E. S., Eisenhart, T. T. and Dempsey, J. L. (2017). A practical beginner's guide to cyclic voltammetry, *Journal of Chemical Education* **95**(2): 197–206.

ExxonMobil (2019). Outlook for energy: A perspective to 2040.

Fischer, F., Sugnaux, M., Savy, C. and Hugenin, G. (2018). Microbial fuel cell stack power to lithium battery stack: Pilot concept for scale up, *Applied Energy* **230**: 1633 – 1644.

Houache, M. S., Hughes, K. and Baranova, E. A. (2019). Study on catalyst selection for electrochemical valorization of glycerol, *Sustainable Energy & Fuels* .

Kim, H. J., Lee, J., Green, S. K., Huber, G. W. and Kim, W. B. (2014). Selective glycerol oxidation by electrocatalytic dehydrogenation, *ChemSusChem* **7**(4): 1051–1056.

Kodera, F., Kuwahara, Y., Nakazawa, A. and Umeda, M. (2007). Electrochemical corrosion of platinum electrode in concentrated sulfuric acid, *Journal of Power Sources* **172**(2): 698–703.

Kwon, Y., Schouten, K. J. P. and Koper, M. T. (2011). Mechanism of the catalytic oxidation of glycerol on polycrystalline gold and platinum electrodes, *ChemCatChem* **3**(7): 1176–1185.

Lee, D., Kim, Y., Kwon, Y., Lee, J., Kim, T.-W., Noh, Y., Kim, W. B., Seo, M. H., Kim, K. and Kim, H. J. (2019). Boosting the electrocatalytic glycerol oxidation performance with highly-dispersed pt nanoclusters loaded on 3d graphene-like microporous carbon, *Applied Catalysis B: Environmental* **245**: 555–568, Appendix A, SI.

Marshall, A. and Haverkamp, R. (2008). Production of hydrogen by the electrochemical reforming of glycerol–water solutions in a pem electrolysis cell, *international journal of hydrogen energy* **33**(17): 4649–4654.

Pourbaix, M., Van Muylder, J. and De Zoubov, N. (1959). Electrochemical properties of the platinum metals, *Platinum Metals Review* **3**(2): 47–53.

Rajalakshmi, N. and Dhathathreyan, K. (2008). Nanostructured platinum catalyst layer prepared by pulsed electrodeposition for use in pem fuel cells, *International Journal of Hydrogen Energy* **33**(20): 5672 – 5677.

Rodrigue, J.-P., Comtois, C. and Slack, B. (2016). *The geography of transport systems*, Routledge.

Simões, M., Baranton, S. and Coutanceau, C. (2012). Electrochemical valorisation of glycerol, *ChemSusChem* **5**(11): 2106–2124.

Tegou, A., Papadimitriou, S., Kokkinidis, G. and Sotiropoulos, S. (2010). A rotating disc electrode study of oxygen reduction at platinised nickel and cobalt coatings, *Journal of Solid State Electrochemistry* **14**(2): 175.

Verma, S., Lu, S. and Kenis, P. J. (2019). Co-electrolysis of co_2 and glycerol as a pathway to carbon chemicals with improved techno-economics due to low electricity consumption, *Nature Energy* p. 1.

Wei, C., Sun, S., Mandler, D., Wang, X., Qiao, S. Z. and Xu, Z. J. (2019). Approaches for measuring the surface areas of metal oxide electrocatalysts for determining their intrinsic electrocatalytic activity, *Chemical Society Reviews* **48**(9): 2518–2534.

Winiwarter, A., Silvioli, L., Scott, S. B., Enemark-Rasmussen, K., Sariç, M., Trimarco, D. B., Vesborg, P. C., Moses, P. G., Stephens, I. E., Seger, B. et al. (2019). Towards an atomistic understanding of electrocatalytic partial hydrocarbon oxidation: propene on palladium, *Energy & Environmental Science* **12**(3): 1055–1067.

Zalineeva, A., Baranton, S. and Coutanceau, C. (2015). How do bi-modified palladium nanoparticles work towards glycerol electrooxidation? an in situ ftir study, *Electrochimica Acta* **176**: 705–717.

Zhang, Z., Xin, L. and Li, W. (2012). Electrocatalytic oxidation of glycerol on pt/c in anion-exchange membrane fuel cell: cogeneration of electricity and valuable chemicals, *Applied Catalysis B: Environmental* **119**: 40–48.

Zhao, Y., Li, X., Schechter, J. M. and Yang, Y. (2016). Revisiting the oxidation peak in the cathodic scan of the cyclic voltammogram of alcohol oxidation on noble metal electrodes, *RSC Advances* **6**(7): 5384–5390.

Metalated Ion-Exchange Resins as Oxygen Selective Adsorbents for Air Separation

Abbie Watson and Remy Shaieb
Department of Chemical Engineering, Imperial College London, U.K.

Abstract

The industrial oxygen market will reach 68bil US$ by 2025 and serves a huge number of industries globally. PSA offers lower capital and energy costs than conventional cryogenic air separation. This makes it more applicable for serving climate change solutions such as oxy-fuel combustion. To serve this purpose an oxygen selective adsorbent presents significant advantages, such as improved efficiency and a lower required uptake. The challenge is to develop an oxygen selective adsorbent with a high selectivity to oxygen and high capacity. Commercially available ion exchange resins have potential to fit this application. Their high loading capacity can be exploited to load the resins with large amounts of metal, which have a high affinity for oxygen, to increase selectivity to oxygen. The aim of this study was to synthesis, characterise and test the validity of these adsorbents. The ion exchange resins used were Amberlite IR-120, a strongly acidic resin with sulfonate functional group, and Amberlite CG-50, a weakly acidic crosslinked methacrylate resin. These were metalated with chromium, cobalt, copper and iron then characterised using TGA, FT-IR and porosity isotherms, and tested. The results show the resins display a low selectivity to oxygen of 0.58-2.66, and an extremely low uptake, achieving 0.08 to 0.51 cm3/g, making them unsuitable for this application.

Introduction

The global industrial oxygen market size will increase to 68bil US$ by 2025, from 43bil US$ in 2018, at a CAGR of 6.8%.[1] The market for oxygen is a large, very fast-growing industry which represents a large number of applications. Oxygen is a common raw material in many industries including metal manufacturing, chemicals and refining, water and waste treatment, energy production, paper and pulp manufacturing.[2] Although oxygen is an abundant gas in the atmosphere, many uses of oxygen require it to be pure and therefore require oxygen separation from air. Due to the much larger proportions of nitrogen in air than oxygen, using air greatly increases the size and therefore cost of equipment. Oxygen can, therefore, be used in oxidation processes to increase efficiency and capacity. An example of this is in steel manufacturing, which is responsible for the largest consumption of oxygen.[1] Oxygen is needed for efficient combustion to occur and reach high thermal efficiencies.[3] Additionally, in the energy generation sector, burning fuels in air produces various other pollutants including NOx and particulates.[4] There is, therefore, a great need to efficiently separate oxygen from air.

There is increasing pressure to reduce greenhouse gas emissions, due to the effects of contributing to climate change. Burning carbon-based fuels produces substantial carbon dioxide emissions, which is a potent greenhouse gas. In the U.S. in 2017, approximately 50% of greenhouse gases emissions are associated with burning fossil fuels for energy and electricity. Carbon capture and storage (CCS) processes can sequester these emissions into storage; however, this requires energy intensive separation to achieve pure CO_2 from the flue gas effluent of CO2, H_2O and pollutants such as NO_x.[5] Oxy-fuel combustion is a promising way to reduce the need for this energy intensive step as the separation occurs pre-burning. Oxy-fuel combustion involves burning the fuel in a high purity stream of oxygen, of which the optimal purity of oxygen is about 95%[11]. This will produce an effluent which is highly pure CO_2 and H_2O, which is easily separated, with high thermal efficiency.[4] This is therefore much more beneficial and can be more economical if a low energy air separation process can be used. This highlights the need for a minimal energy intensive method for air separation to achieve high purity oxygen.

There are currently 3 main methods of oxygen separation from air. Cryogenic separation is the oldest and most commonly used method of oxygen separation, developed in the late 19th century to produce high purity oxygen in large volumes. This is where air is cooled to separate components based on boiling points and thus air must be cooled to below -192 ºC, at 1 bar.[6] This process is favoured if you desire high purity oxygen (>99.5%) at high volumes, with the current largest single train plant producing 5250 tonnes/day.[6] Advantages include that a high purity nitrogen stream is produced also, which can be commercialised and has low electricity use per unit nitrogen. Disadvantages include large space and utility requirement, high capital cost and long start up and shutdown.[9] Even though cryogenic separation of air is the current leading method, its major flaw is the large energy requirement.[3] This makes it unsuitable as an application to produce oxygen for oxy-fuel combustion because the large energy requirement necessary would potentially off-set the economic and energy potential savings of not requiring downstream separation of the CO_2 effluent.

Another method for air separation is membrane separation. This is an emerging, low energy consuming technology. These membranes consist of bundles of selectively permeable hollow fibres, which allow for selective permeation of oxygen and other gases based on diffusivity and solubility.[7] Membrane technology can produce 10 to 25 tonnes per day with an oxygen purity of 25 to 40%.[8] Membranes, therefore, have specific applications where low purity and low volumes are required, whilst having a flexible output, low cost and easy installation.[9] This can be useful for some medical applications and increasing thermal efficiency in the combustion of natural gas.[10] Membranes, however, cannot be used to produce oxygen suitable for oxy-fuel combustion, due to the low purity and low volumes of oxygen produced. For oxy-fuel combustion the required purity is in the range 90-95%. Additionally, a thermal

power plant of 1 GWe electrical power would require about 18 000 tons of pure oxygen per day.[11]

This report will focus on air separation via adsorption. Pressure swing adsorption is a method where gases are separated from a mixture based on having a varying affinity for an adsorbent material, typically a molecular sieve or other porous material with a high surface area. The adsorbents are selective, and will adsorb one species, the target, over the others. This can be due to polarizability, as with zeolites, adsorption equilibrium or a difference in sorption rates, as with molecular sieves.[12] Adsorption is widely used industrially today, often for medical applications and paper and pulp manufacturing. It has a production capacity of up to 300 tonnes per day of relatively high purity oxygen (~95%),[13] with the impurity almost entirely argon which is very difficult to separate from oxygen.[12] PSA is, therefore, be applicable for natural gas plants of capacity < 25 MWe.[11] Consequently, this process can be more economical than cryogenic separation for low to medium production needs and benefits from relatively low to medium capital costs. The major advantage is the feature of operating at ambient conditions, which saves energy compared to the temperature requirements of ion transport membranes and cryogenic cooling. The major drawback is the low selectivity values of O_2 and N_2.[3] Other disadvantages include noise, high maintenance requirements and limited scalability.[9]

Current adsorbents include zeolites, activated carbon silica gel and alumina.[3] These all have very large specific surface areas, up to the region of 1000 m^2/g[14] which is crucial so many gas molecules can be adsorbed, as they are attracted to the surface. The most common adsorbents, zeolites, are natural and synthetic minerals possessing micro-porous silica-alumina crystalline structures that have shape selective properties. Zeolites more strongly adsorb nitrogen since it is more polarizable than oxygen. These have a selectivity towards nitrogen of around 2 to 6 at 1 bar in literature.[3] Zeolites currently achieve nitrogen uptakes within a range of 10-25 cm^3/g, with the most favourable at the higher end.[3,15] By comparison, carbon molecular sieves preferentially adsorb oxygen due to the smaller size of oxygen entering the pores more easily.[12] In research these have achieved selectivity of O_2: N_2 of up to 3.58 and uptakes of oxygen up to 6.7 cm^3/g, at 1atm and 0°C.[16]

Current adsorption processes for air separation revolve around adsorbing nitrogen from air. However, an oxygen selective adsorbent would lead to cost and efficiency benefits since oxygen is the less abundant component. Theoretically, 4x as little adsorbent would be needed.[17] In addition, the remaining oxygen is difficult to separate from argon in the air under normal conditions by most zeolite adsorbents. Oxygen selective adsorbents would bypass this issue.[18]Adsorption via an oxygen adsorbent, therefore, has the potential to be a highly energy efficient method to produce higher purity oxygen, free from argon.

Current oxygen adsorbents, such as carbon molecular sieves, feature lower than desired selectivity's and uptake capacity, making them unfeasible for industrial applications. Thus, the challenge is to develop an oxygen selective adsorbent with high selectivity to oxygen that can adsorb a large quantity of oxygen.

This paper will investigate the suitability of ion exchange resins loaded with different metals as adsorbents for oxygen separation from air. This is due to the knowledge that oxygen adsorbs strongly onto some metals, and significant improvement in N_2/O_2 selectivity was observed by ion exchange of Na+ with Li+ in a LSX zeolite.[3] Additionally, metal ion salts have seen to have selectivity's of O_2/N_2 of 2-3 in current research.[19] Therefore, a resin impregnated with a very high metal content, may result in a significant improvement in the adsorption of oxygen onto the surface.

The specific resins used in this study have been chosen for their large ion exchange capacity and are commercially typically used for ion exchange applications such as removing metal pollutants from water streams,[20] and often for water softening. Their high capacity provides the potential for a very high uptake of metal, increasing the likelihood for oxygen adsorption. The resins were also chosen based on their chemistry, with one resin being a strongly acidic polystyrene–divinylbenzene copolymer and the other a weakly acidic crosslinked methacrylate. This different chemistry may offer different results and additional insight.

Should these resins prove effective adsorbents, this process is particularly appealing due to benefiting from the well-established market for these resins which already exists. Firstly, this has the benefit that it is known that they can be effectively brought to market, sold and distributed. This also provides the benefit that there is widely available information on the classification and characterisation of these resins, for example, structure, suitable temperature and pH ranges.[20] They are commercially available in many countries and therefore provide an easily accessible adsorbent that can potentially be quickly synthesised into a functioning adsorbent without the requirement for complex and protracted synthesis.

This paper aims to determine the validity of using these ion exchange resins as an adsorbent for air separation by pressure swing adsorption. Firstly, the synthesis of adsorbents from industrial ion exchange resins will be investigated, loading with metals. It is desired to synthesise an adsorbent which features a high oxygen uptake capacity, high selectivity to oxygen alongside being thermally stable and resistant to capacity loss. Furthermore, the resins will be characterised based on metal content, bond structure, thermal stability and surface area. Finally, the metalated resins will be tested by adsorbing N_2 and O_2 at ambient conditions to determine how successful an adsorbent for air separation they are. The resins will be compared to values which constitute current success of oxygen selective adsorbents in the literature to determine their validity as effective adsorbents. Based on the data discussed thus far this has been determined as a selectivity to oxygen of 3, and an oxygen uptake of 6 cm^3/g.

Background

Patents on the subject of pressure swing adsorption were filed from 1930 onwards.[21] PSA methods of separating gaseous mixtures are usually credited with being introduced by Skarstrom[22] and Guerin de Montgareuil and Domine[23] in 1957-58. While these were not the first patents filed regarding PSA, Guerin de Montgareuil and Domine introduced the use of a vacuum swing, and Skarstrom used a low-pressure swing to purge the adsorbent bed and regenerate the absorbent, developing what we know today as PSA. Kahle was the first to publish papers outlining the principle and details of PSA for removal of CO_2, water and hydrocarbons from air, in 1953.[24] Zeolites began being produced commercially in 1955 and oxygen selective carbon molecular sieves (CMS) in 1972.[12]

Throughout these developments, there was a delay before the technology was taken up. PSA did not gain widespread commercial acceptance until the 1980s when the number of publications and patents with the word PSA soared. The reason for this may be due to opposing interests in the cryogenic separation industry and lack of familiarity with the underlying principles.[12]

Air separation was very quickly recognised as a use for pressure swing adsorption, it was one of the applications described by Skarstrom in his early patents.[22] The applied development of air separation technologies increased in the 1980s alongside the PSA technology itself. Nitrogen selective zeolites were applied for air separation and Ruthven comments in his 1994 print that zeolites 5A and 13X are commonly used adsorbents with selectivity's to nitrogen of 3-3.5.[12] Modern zeolites have a selectivity of nitrogen to oxygen of up to 6.[3]

Development of zeolites and other adsorbents continued throughout the 90s. Zeolites became very popular adsorbents for a variety of industries and maintained the focus of a lot of research.[3] The opportunity for improvement remained with increasing selectivity and developing stable O_2 selective adsorbents. Gaffney commented in 1996 that the last few years had seen development of O_2 adsorbents with high selectivity. Additionally, CMS's were developed for the selective adsorption of oxygen based on the principle of oxygen diffusing more rapidly through molecular pores due to its smaller size. Developments of CMS's continued and in 1993 Golden investigated the use of O_2 selective inorganic salts dispersed in the macropores of CMS's. This was an early example of the use of metal to increase oxygen selectivity and achieved an increase of 2-4% nitrogen recovery.[25]

Investigating the use of ion exchange resins as adsorbents emerged in the 70s. Abrams investigated their use for the removal of organic contaminants from aqueous solutions in 1974.[27] Similar, Kunin investigated polymeric adsorbents for waste effluent treatment.[26] In Ruthven's 1994 Pressure Swing adsorption, a classification of adsorbents is presented, where "Polymeric resins containing an -OH group or cations" is specified.[12]

It is, therefore, clear that ion exchange polymeric resins were being investigated in the context of adsorption, however, what classifies an adsorbent as successful needs to be understood.

Abrams reasoned that adsorbents should possess some chemical functionality, be sufficiently porous to allow easy diffusion into and out of the resin matrix and possess a high enough concentration of polar groups to facilitate chemisorption on the surface.[27] More generally, in a 2019 review of using adsorbents for CO_2 capture, the following criteria were described for an adsorbent: Adsorption capacity for adsorbate, selectivity for adsorbate, mechanical strength of particles, chemical stability, regeneration of sorbents.[28]

Grande presents a thorough review of advances in PSA in his 2012 publication. He describes how the development of materials science in the last half century has resulted in the discovery of many potential adsorbents. However, an easily regenerated adsorbent is required. It is common to find high capacity adsorbents; however, these adsorbents often exhibit isotherms which are very steep, meaning a large vacuum is required for desorption, which is energy intensive and defeats the purpose PSA fulfils of a low energy alternative method. So, materials showing linear or slightly nonlinear isotherms are preferred in PSA design.[29] This highlights the need for an adsorbent facilitating a low energy process.

Clearly, there is a large variety factors which determine the success of adsorbents and the use of said adsorbent will be a determinant. In the ion exchange resin context, an adsorbent with a high concentration of the functional group, mobility in out and of the resin which is easily regenerated appear to be the core factors. Specifically for gas separation, the common factor mentioned is the necessity for an oxygen selective adsorbent with a high selectivity of oxygen to nitrogen.

Recently, the use of these resins has diversified. A patent for the removal of acid gases from gaseous mixtures using ion exchange resins was filed in 1997.[30] In 2010, a patent was filed outlining the process and apparatus to use ion exchange resins as adsorbents in CO_2 capture.[31] This demonstrates the development in the understanding of ion exchange resins uses and testament that they could be equally successful in the context of air separation.

A 2018 review in Progress in Energy and Combustion Science regarding the application of O_2 separation for oxy-fuel combustion highlights the need for oxy-fuel combustion, explains the options for oxygen separation methods and analyses the historical and recent advancements. It is a thorough review of the options, opportunities and challenges for adsorption as a method to make carbon capture and sequestration a feasible solution to tackle rising greenhouse gases.

In summary, there has been a long development of PSA, yet there are still more developments to be made. In particular, an oxygen selective adsorbent with high selectivity, that can be produced easily and generate large quantities of oxygen at ambient conditions is the ideal process for an oxy-fuel application.

Methods

Material Synthesis

The ion exchange resins used were Amberlite IR-120, a strongly acidic polystyrene–divinylbenzene copolymer with a sulfonate function group[32] and Amberlite CG-50, a weakly acidic macroporous crosslinked methacrylate with a carboxylic acid functional group.[33]

Firstly, 5 grams of each resin were washed 10 times with 1M HCl and rinsed with deionised water. The resins were then soaked overnight in saturated solutions of the following metals : Chromium (III) nitrate, Cobalt (II) nitrate, Copper (II) nitrate and Iron (III) chloride. Subsequently, they were washed with deionised water and acetone and dried.

Material Characterisation

The samples were characterized using Fourier transform infrared spectroscopy (FT-IR) spectroscopic, thermogravimetric techniques and porosity measurements.

For the FT-IR measurement, the samples were first ground to a powder using an agate mortar. Subsequently, the sample was loaded onto the machine and tested in the wavenumber range of 400−4000 cm^{-1} using an Agilent Cary FT-IR spectrometer. The resulting output was the IR spectra of the resins.

The metal content and thermal stability were determined by measurement in a thermogravimetric analyser TA TGA 500 and a Netzsch TG209 F1 Libra thermogravimetric analyser. Prior to analysis, the samples were vacuumed at 120°C overnight. The samples were then loaded into a platinum crucible and subsequently heated under an air atmosphere (flow rate of 50 $mL.min^{-1}$ with protective N_2 at 30 $mL.min^{-1}$) from 30 to 900 °C on a ramp rate of 5°$C.min^{-1}$. An isothermal step was used at 120 °C for 150 minutes to ensure full removal of moisture.

The surface area was measured using a porosity and surface area analyzer (Quantachrome iQ3) at 77 K using liquid nitrogen. The samples were degassed in situ for 12 h at 120 °C. The surface areas of the samples were determined using the Brunauer−Emmett−Teller (BET) method[34]. The total pore volume was ascertained from the volume of N_2 adsorbed at approximately 0.98 bar relative pressure (P/P0).

Material testing

Adsorption and desorption isotherms were measured using a Micromeritics 3Flex sorption analyzer up to 1 bar at 25 °C using a water bath to control the temperature. The samples were degassed overnight (120 °C, ~0.2 mbar) and further degassed in situ for 6 h (0.003 mbar) before the analysis.

Table 1: Metal uptake of all resins on mass, molar and % uptake basis

Sample	Oxide	Mass of metal per gram of dry sample (mg/g)	Mass of metal per gram of sample (mmol/g)	% uptake compared to expected capacity
Cr IR-120	Cr_2O_3	79	1.52	203%
Co IR-120	Co_3O_4	376	6.38	567%
Cu IR-120	CuO	6	0.10	9%
Fe IR-120	Fe_2O_3	86	1.54	137%
Cr CG-50	Co_3O_4	27	0.52	16%
Co CG-50	Cr_2O_3	12	0.2	4%
Cu CG-50	CuO	9	0.15	3%
Fe CG-50	Fe_2O_3	51	0.91	18%

Results and Discussion

Metal Uptake

After metalation of the ion exchange resins, Thermal Gravimetric Analysis (TGA) was carried out to find the mass of metal contained within the resins. Metal oxides have thermal stabilities in air of greater than 1000°C[35], whilst ion exchange resins usually have thermal stabilities up to 150°C (initial onset temperature)[36]. Thus, the resins were heated to 900°C, to leave the oxide of the metal behind. This was then weighed to calculate the mass of metal originally in the resin.

Metal loading is crucial characterisation as it allows us to discover whether the metal sites are what is leading to oxygen uptake and selectivity, or whether the oxygen absorption is physisorption only. Metal loading was purported to improve the selectivity to oxygen.[37]

Amberlite IR-120 has a lower *nominal* capacity than CG-50, at 2.25 mmol/g[38] and 10 mmol/g[39] respectively. Thus, one could expect the metal uptake to be higher for CG-50 due to higher capacity and because the particle size is smaller. Smaller particle sizes lead to less resistance to intra-particle diffusion, which becomes limiting at higher concentrations of metal such as in the saturated solutions used.

However, table 1 shows that the uptake was much lower for Amberlite CG-50. The uptake of Cobalt is over 30 times greater in IR-120 than CG-50, at 376 mg/g and 12 mg/g respectively. It is postulated that this is because Amberlite CG-50 is a weakly acidic exchange resin, meaning its uptake capacity is severely hindered by acidity in the target solution.[40] The solutions used were saturated solutions of the metal ions, which have a polarising effect on water, leading to acidic solutions. In general, weakly acidic exchange resins are not used in solutions lower than pH 6[41]. Due to the high metal content of each saturated solution, the pH values of our solutions were likely greater than this.

The pK_a values for the metal ions used were 3.7, 9.7, 7.5 and 2.2 for Chromium (III) nitrate, Cobalt (II) nitrate, Copper (II) nitrate and Iron (III) chloride respectively.[42] For comparison, carboxylic acids have pK_a of approximately 5.[43]

To increase the uptake capacity of weakly acidic cation resins, consideration will need to be given to the pH of the prepared solution. Due to the influence on pH of metal ion, the saturated solution may not be the optimum concentration of solution to give maximum uptake. Measurements should be taken to find the optimum concentration. Buffering is also an option to regulate pH. Pre-swelling the beads using ethanol or other solvent is also a possible method of increasing uptake.[44]

Table 1 shows that the loading for Amberlite IR-120 exceeds the nominal capacities, with uptakes of 137, 203 and 567%. We can be assured that when testing the air separation properties of the IR-120 resins, we are working with a metal uptake that is expected of the resin. Copper is the exception; however, the uptake of copper is low in both resins so this may be due to a differing property of the copper ion. This should be explored in more detail.

When comparing the metal uptake for the different metals, we find no link between the hydrated radii of metals in the solution and the uptake. The hydrated radii are 4.61, 4.23, 4.19 and 4.28 Å for Cr^{3+}, Co^{2+}, Cu^{2+} and Fe^{2+}. Size selectivity is not the deciding factor here. Chromium has the largest radius yet not the lowest uptake. Nor is there a link with quantity of metal dissolved to form the saturated solution. Therefore, it is possible that the differences in uptake between metals are due to disparities in the preparation procedure of each
metal, including degree of stirring and other inhomogeneities.

Repeat TGA readings should be taken, along with other experiments to measure uptake such as direct weighing, and x-ray fluorescence.

FT-IR

Taking FTIR spectra is important as it allows us to search for traces of metal-function group interactions to ensure that the metal we have taken up is located at the resin active sites, and not physically trapped in the structure. This ensures that when the resin is used industrially, the metal will remain at the active sites – ensuring minimal loss of uptake capacity over time. FTIR signals also give validation to the structures of the resins.

Figure 1 shows the IR spectra of the pure and metallated Amberlite IR-120 and CG-50 resins respectively.

For Amberlite IR-120 each sample shows a broad peak around 3400 cm^{-1}, due to water in the sample. Peaks at 2900 are due to the C-H stretching modes on the methyl groups and on the aromatic ring. The strong band at 1600 cm^{-1} is due to the C=C bond in the aromatic ring. The very strong bands observed in the infrared spectrum near 760cm^{-1} and 690cm^{-1} confirm the presence of a monosubstituted aromatic group. These bands are due to the CH out-of-plane vibration and a ring out-of-plane deformation respectively.[45]

The sulfonate groups exhibit peaks at 1160, 1345, 2350, 2760 cm^{-1}.[46] At 1250 cm^{-1} we see new peaks which are present in the metallated spectra but not in the pure spectra. These represent the formation of sulfonic acid metal salts.[46] This proves that metallation has occurred at the active sites.

For Amberlite CG-50 resins, the sharp peak at 1740 cm^{-1} is due to the C=O stretching vibration of the carboxylic acid groups[47]. The broad peaks at 2900 cm^{-1} are due to the OH bondson the carboxylic acid.[48] This confirms the carboxylic acid group as the functional group for ion exchange in Amberlite CG-50. Thus we would expect spectral signals for carboxylic acid salts to be present in the metallated resins but not in the pure resin. Carboxylic acid salts exhibit signals in the region 1695 cm^{-1} – 1540 cm^{-1} range. [49] Fe IR-120 and Cu IR-120 show strong signals in this region, whereas the other metallated resins show minor distortions compared to the pure resin. This may be due to the low metal uptake of the CG-50 resins in general.

FTIR analysis has confirmed the structure of both resins including functional group and polymer backbone. Metallation at functional sites has been strongly confirmed in IR-120 but only weakly in CG-50. More tests should be done to confirm that the uptake is at the functional group in CG-50.

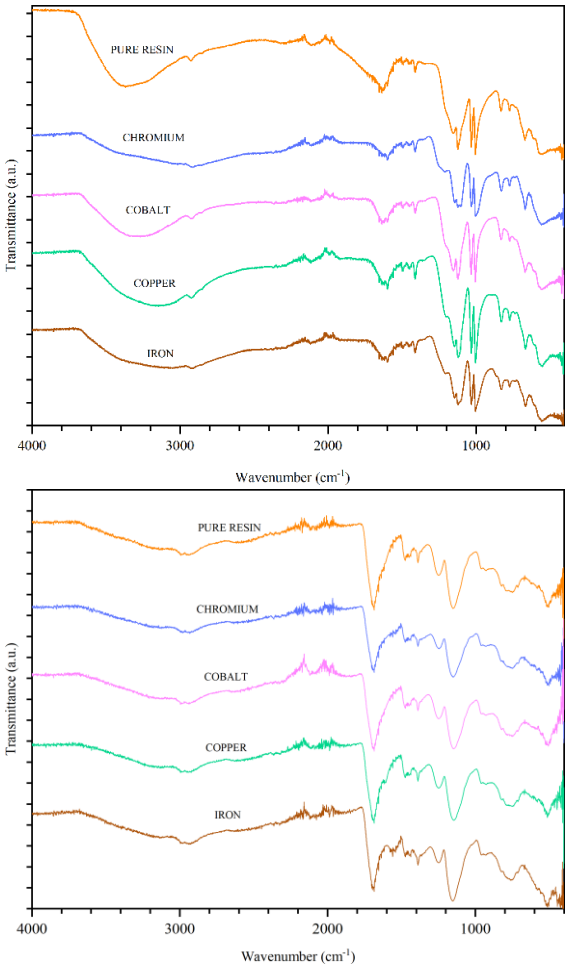


Figure 1. FTIR Spectra of Amberlite IR-120 and Amberlite CG-50 respectively

Thermal stability

Thermal stability is an important characterisation to make as the resins may be used under high temperature in industrial gas separations. Any breakdown of the structure may lead to a loss of metal or of the functional acidic groups which coordinate the metals, reducing oxygen uptake ability. Furthermore, the loss of physical properties may result in an increase in the pressure drop across a bed of resin, increasing operating costs.[50]

Figure 2 shoes the TGA curves for the two resins. For Amberlite IR-120 the pure resin shows two distinct regions of mass loss. One commencing at 225°C and the other at 500°C. This

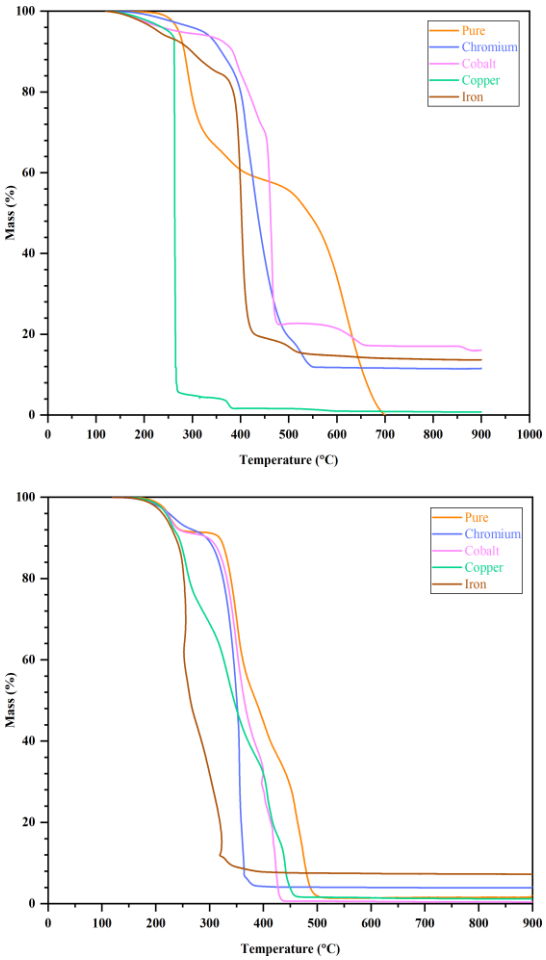


Figure 2. TGA curves for Amberlite IR-120 and Amberlite CG-50 respectively

distinction could be explained by the loss of the sulfonate groups occurring first, followed by the poly(styrene) backbone which requires more energy to decompose. This is consistent with literature on cation exchange resins. Stalbush et al. remark that when cation exchange resins are heated desulfonation of the resin occurs and inorganic sulfate is released from the resin. Above temperatures of 400-500 °C rapid degradation of the copolymer backbone occurs, resulting in carbonization of the resin and the formation of carbon dioxide, sulfur oxides and water when heated in the presence of air or oxygen.[50]

The metallated resins do not show two distinct regions - the slopes are monotonic. The Initial Decomposition Temperature (IDT) is similar for metallated and pure resins, however the maximum rate decomposition temperature (MRDT) is higher in the metallated resins. This may be due to coordination of the metal groups increasing stability of the sulfonate group, in contrast to the protonated sulfonate groups that occur in the pure resin. Chelation is known to increase stability in polymer resins. The multivalent atoms act as multifunctional crosslinking agents.[51]

Copper appears to be the anomaly. This is likely due to experimental procedure, as the rate of decomposition is very high, suggesting possible bumping of the resin beads.

For Amberlite CG-50, two distinct regions of mass loss are seen for the pure sample, along with chromium and cobalt. Copper and iron show one monotonic mass loss. This suggests that chromium and cobalt have been altered by metalation in a way that does not influence the thermal stability, whereas copper and iron have been altered in a way that does influence the thermal stability.

The two regions could correspond to the removal of the functional carboxylic acid group and the decomposition of the methacrylate backbone. However, there is little difference between the samples overall, with all IDTs being the same and all MRDTs occurring between 300°C and 400°C. This is possibly because the metal doesn't influence stability, or the metal loading is too low to give a difference.

Both pure resins show a complete decomposition. Pure IR-120 has decomposed fully at 700°C, and pure CG-50 has decomposed fully by 500°C. This adds credibility to the metal loading masses as they are relative to a mass of zero.

Surface area

Surface area measurements were carried out for all samples in order to discover which samples were the most promising to take forward to the testing phase of the project. Surface area gives a reasonable estimation of the oxygen uptake potential of the resins, since high porosity means good access to the metal sites.

Table 2: Porosity data for Amberlite resins IR-120 and CG-50 loaded with metals derived using N2 at 77K on a Quantachrome iQ3.

	Surface area (m^2/g)	Max volume adsorbed (cm^3/g)
IR-120	6	0.36
Cr IR-120	6	0.37
Co IR-120	10	0.56
Cu IR-120	10	0.63
Fe IR-120	7	0.41
CG-50	6	0.48
Cr CG-50	6	0.42
Co CG-50	13	0.24
Cu CG-50	3	0.21
Fe CG-50	5	0.39

For both resins the isotherms show few micropores, as seen in figure 3 by the initial small slope. This is unpromising as it means less access to the metal sites. Positively though, in general the roughly linear profiles mean that the cycling capacity is a high proportion of the overall gas uptake since strong vacuums don't need to be pulled to regenerate the adsorbents.

For Amberlite IR-120 the surface area for the pure resin was not higher than for the metallated resins – metallation does not decrease surface area. For Amberlite CG-50 the surface area for the pure resin was

higher than for the metallated resins – so metallation decreases surface area. This may be as metal is blocking the surface of the resin, preventing acess to the internal pore space.

The surface areas are all within a relatively narrow range. Random error or any inhomogeneities may have caused data to exhibit noise. It is not due to machine insensitivy as the Micromeritics 3Flex sorption analyzer measures to a minimum of 0.01 m^2/g[52].

The similarities in surface areas mean that neither Amberlite IR-120 nor CG-50 stood out as being more promising. Thus surface area was not used to distinguish samples when moving to the testing phase of the study.

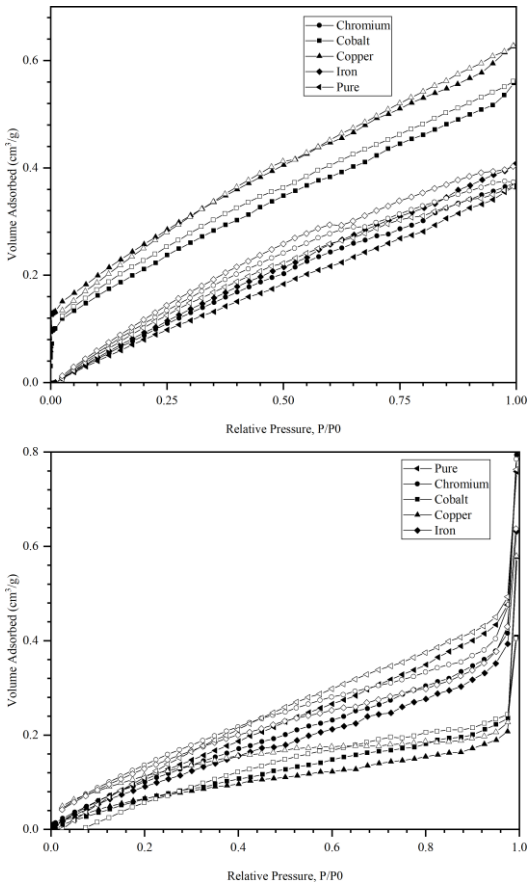


Figure 3. Isotherms for Amberlite IR-120 and CG-50 respectively, with N_2 at 77K

Air separation

After characterisation, the most promising resins were chosen, and adsorptions with oxygen and nitrogen at room temperature and pressure were carried out. Overall uptake and selectivity to oxygen were calculated.

Based on the characterisations, Amberlite IR-120 was chosen as the most promising resin due to high uptake of metal. The pure resin and three metallated samples were tested, discarding copper. Pure CG-50 and Fe CG-50 were tested for comparison.

Table 3: Selectivity and uptake data of the resins testing with both O2 and N2 at 298K.

The selectivities were in the range 0.58 to 2.66 and the oxygen uptake in the range 0.08 to 0.51 cm^3/g. These are lower than the minimum criteria for success of 3 and 6 for selectivity and oxygen uptake respectively. Adsorbents in industry are one order of magnitude higher in uptake, and have moderately higher selectivity.

Table 3 shows that the oxygen and nitrogen uptake (at 1 bar) for IR-120 resins is lower than the uptake of the pure sample. This either means that the data is erroneus and thus is dominated by noise, or that the metal loading is resulting in lower gas adsorption per gram of sample. Metal could be blocking the surface of the resins, preventing access to the internal pore space.

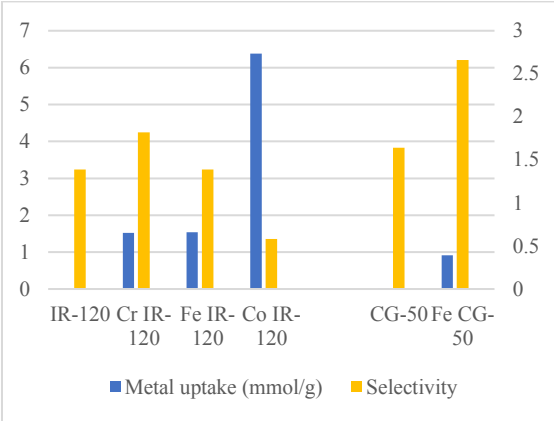


Figure 4. Chart of metal uptake (mmol/g) and selectivity towards oxygen

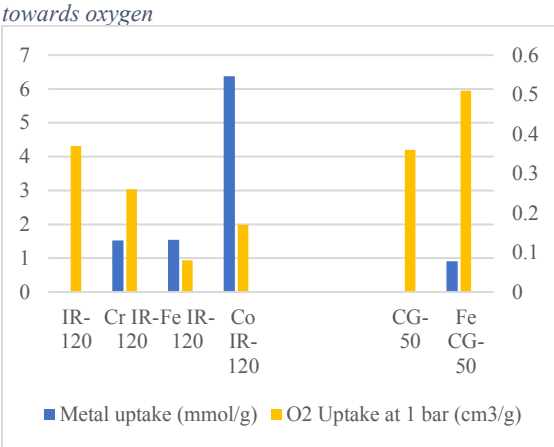


Figure 5. Chart of metal uptake (mmol/g) and oxygen uptake at 1 bar (cm^3/g)

Figure 4 shows selectivity plotted with metal uptake. It shows that for IR-120, metal uptake doesn't appear to correlate with selectivity.

For CG-50 the metal uptake gives a considerable increase in selectivity (going from 1.7 to 2.7), despite having a lower metal uptake than the IR-120 resins. More testing should be done to truly assess the feasibility of CG-50 resins.

		IR-120	Cr IR-120	Co IR-120	Fe IR-120	CG-50	Fe CG-50
Uptake at 1 bar (cm^3/g)	O2	0.37	0.26	0.17	0.08	0.36	0.51
	N2	0.26	0.14	0.12	0.13	0.22	0.19
Selectivity of O2:N2	At 1 bar	1.39	1.82	1.39	0.58	1.64	2.66
	At partial	0.76	1.14	1.05	0.30	0.53	0.59

Conclusion

This study aimed to synthesise an oxygen selective adsorbent that features a high oxygen uptake capacity, high selectivity to oxygen alongside being thermally stable and resistant to capacity loss. These parameters were characterised via the metal uptake, bond structure, thermal stability, surface area and finally testing gave oxygen to nitrogen selectivity.

It can be evaluated based on the evidence presented in this study that metalated, commercially available, ion exchange resins are not suitable adsorbents for air separation via PSA.

Firstly, the metal uptake was evaluated. A high uptake was seen for three resins of 137% for iron, 203% for cobalt, 567% for chromium. However, overall this was much lower than desired as five out of eight metalations achieved just 4-18% of the nominal capacity uptake. We have speculated reasons as to what may have caused these trends. The poor metalation seen is likely to be due to preparation techniques. The preparation used in this study may be insufficient to facilitate the desired metalation. For example, the solutions may not have been effectively saturated or there was insufficient mixing in the solutions to ensure metalation throughout the whole structure. In future, additional measures should be taken to improve this. The metal solutions used could be measured to ensure saturation or one could improve the mixing procedure by allocating more space to samples. Finally, repeat measurements and other experiments can be done to validify results, such as direct weighting or x-ray fluorescence.

The thermal stability was determined as 225°C for both resins. In the case of Amberlite IR-120, an increase in thermal stability was seen through metalation of the resin, manifested as one region of decomposition compared to two for the pure un-metalated resin, as the metal coordination likely increases stability of the sulfonate group.

For Amberlite CG-50, there is little difference between pure and metalated thermal stability. Therefore, either metal doesn't influence stability, or the metal loading is too low to impact a significant difference.

The resins display a low selectivity to oxygen of 0.58-2.66, not meeting the target selectivity of 3. Secondly, the uptake was extremely low achieving 0.08 to 0.51 cm^3/g, a very small amount relative to our target of 6 cm^3/g.

A possible reason for these results is that there may be insufficient access to the metal sites due to low porosity which leads to a low uptake of oxygen. Additionally, metalation may not be an effective means to increase selectivity in the measure that we anticipated; this would result in lower selectivity than was aimed for. Alternatively, the metalation could not sufficient enough to see a significant change in selectivity.

One could investigate methods to increase the oxygen uptake of the resins. Methods such as pre-swelling of resins may be employed using ethanol to increase the pore size. Alternatively, one could increase pH or buffer the solution to increase uptake for weakly acidic resins, such as the CG-50 resin used. Furthermore, an alternative starting material could be used. One could investigate the use of porous metalated polymers, as the porous structure would provide access to the metal sites to facilitate the oxygen uptake.

Acknowledgments

We would like to thank Robert Woodward and David Danaci for their outstanding help and guidance and throughout this study. We are particularly grateful for their support during the experimental aspects of the work

[1] MarketWatch. (2019). Industrial Oxygen Market 2019 Global Industry Share, Demand, Top Players, Industry Size, Future Growth by 2025: Research Reports World. [online] Available at: https://www.marketwatch.com/press-release/industrial-oxygen-market-2019-global-industry-share-demand-top-players-industry-size-future-growth-by-2025-research-reports-world-2019-03-27 [Accessed 16 Nov. 2019].

[2] Praxair.co.in. (2019). Uses and Applications of Oxygen Gas and Liquid Oxygen by Industry. [online] Available at: http://www.praxair.co.in/gases/buy-liquid-oxygen-or-compressed-oxygen-gas?tab=industries [Accessed 16 Nov. 2019].

[3] Wu, F., Argyle, M., Dellenback, P. and Fan, M. (2018). Progress in O2 separation for oxy-fuel combustion–A promising way for cost-effective CO2 capture: A review. Progress in Energy and Combustion Science, 67, pp.188-205.

[4] Rao P. and Muller M. (2007) 'Industrial Oxygen: Its Generation and Use', *ACEEE Summer Study on Energy Efficiency in Industry*, **6**, 124-126. Available at: https://aceee.org/files/proceedings/2007/data/papers/78_6_080.pdf [Accessed 15 Nov. 2019]

[5] Kumar, D. and Kumar, D. (2018). Sustainable management of coal preparation. 1st ed. Woodhead Publishing, pp.15-33.

[6] Engineering Division, The Linde Group. (2019) History and technological progress: Cryogenic air separation. Pullach, Germany. Available at: https://www.leamericas.com/en/images/Cryogenic%20air%20separation%20brochure19_4353_tcm136-414865.pdf [Accessed 15 Nov. 2019]

[7] Chong, K.C.; Lai, S. O.; Thiam, H. S.; Teoh, H. C and Heng, S. L 2016, 'Recent Progress of Oxygen/Nitrogen Separation using Membrane Technology', *Journal of Engineering Science and Technology*, 11, 7, pp. 1016-1030. Available at: http://jestec.taylors.edu.my/Vol%2011%20issue%207%20July%202016/11_7_8.pdf [Accessed 14 Nov. 2019]

[8] Sanders, D.F.; Smith, Z.P.; Guo, R.; Robeson, L.M.; McGrath, J.E.; Paul, D.R.; and Freeman, B.D. (2013). Energy-efficient polymeric gas separation membranes for a sustainable future: A review. Polymer, 54, 4729-4761.

[9] Aprilia Jaya, KLM Technology Group. (2013) AIR SEPARATION UNITS (ENGINEERING DESIGN

GUIDELINE). Version 1. Johor Bahru, Malaysia. Available at: http://kolmetz.com/pdf/EDG/ENGINEERING_DESIGN_GUIDELINE_Air_Separation_Units_Rev01web.pdf [Accessed 15 Nov. 2019]
[10] Nagar, H.; Vadthya, P.; Prasad, N. S.; & Sridhar, S. (2015). 'Air separation by facilitated transport of oxygen through a Pebax membrane incorporated with a cobalt complex', RSC Advances, 5, 76190-76201.
[11] Banaszkiewicz, T.; Chorowski, M. and Gizicki, M. **(2014).** "Comparative analysis of cryogenic and PTSA technologies for systems of oxygen Production" AIP Conference Proceedings **1573**, 1373. https://doi.org/10.1063/1.4860866
[12] Ruthven, D. M..; Farooq, S. and Knaebel, K. S. Pressure Swing Adsorption. New York: VCH, 1994. Print.
[13] Smith, A.R.; and Klosek, J. (2001). A review of air separation technologies and their integration with energy conversion processes. Fuel Processing Technology, 70, 115-134.
[14] Cid, R. (1983). Zeolites surface area calculation from nitrogen adsorption data, *Journal of Catalysis*, **80** (1), pp. 228-230.
[15] Yang, R, T.; Chen, Y, D.; Peck, J, D. and Chen, N. (1996). Zeolites Containing Mixed Cations for Air Separation by Weak Chemisorption-Assisted Adsorption, Industrial & Engineering Chemistry Research, 35 (9), 3093-3099. DOI: 10.1021/ie950783a
[16] Jayaraman, A. and Yang, R. (2005). Stable oxygen-selective sorbents for air separation. Chemical Engineering Science, 60(3), pp.625-634.
[17] Jayaraman, A., Yang, R.T., Cho, SH. et al. Adsorption (2002) 8: 271. https://doi.org/10.1023/A:1021529328878
[18] Xu, M., Wu, H., Lin, Y. and Deng, S. (2018). Simulation and optimization of pressure swing adsorption process for high-temperature air separation by perovskite sorbents. Chemical Engineering Journal, 354, pp.62-74.
[19] Jayaraman, A. and Yang, R. (2005). Stable oxygen-selective sorbents for air separation. Chemical Engineering Science, 60(3), pp.625-634.
[20] Alguacil, F., Alonso, M. and Lozano, L. (2004). Chromium (III) recovery from waste acid solution by ion exchange processing using Amberlite IR-120 resin: batch and continuous ion exchange modelling. Chemosphere, 57(8), pp.789-793.
[21] Finlayson, D. and Sharp, A. J. to British Celanese Corp (1930). UK365092.
[22] Skarstrom, C. W. to Esson Research and Engineering Company. (1958). Method and Apparatus for Fractionating Gaseous Mixtures by Adsorption. 2944627.
[23] Guerin de Montgareuil, P. and Domine, D. to Air Liquide SA. (1957). Process for Separating a Binary Gaseous mixture by adsorption. FR752290A.
[24] Kahle, H. (1953). 'Die reversible Adsorption Als Mittel Zur Vorreinigung Und Zerlegung Von Gasgemischen', *Chemie Ingenieur Technik*, **25**, 3, 144-48. Web. Available at: https://onlinelibrary.wiley.com/doi/10.1002/cite.330250310 [Accessed 18 Nov. 2019].
[25] Gaffney, T. (1996). Porous solids for air separation. Current Opinion in Solid State and Materials Science, 1(1), pp.69-75.
[26] Kunin, R. (1977). Polymeric adsorbents for treatment of waste effluents. Polymer Engineering and Science, 17(1), pp.58-62.
[27] Abrams, I. (1975). Macroporous Condensate Resins as Adsorbents. Industrial & Engineering Chemistry Product Research and Development, 14(2), pp.108-112.
[28] Samanta, A., Zhao, A., Shimizu, G.K.H., Sarkar, P. and Gupta, R., 2012. Post-combustion CO_2 capture using solid sorbents: A review. Industrial and Engineering Chemistry Research, 51(4), pp. 1438-1463.
[29] Grande, C. (2012). Advances in Pressure Swing Adsorption for Gas Separation. ISRN Chemical Engineering, 2012, pp.1-13.
[30] Quinn, R. to National Institute for Strategic Technology Acquisition and Commercialization. (1997). Removal of acid gases from gas mixtures using ion exchange resins. US5797979A.
[31] Tirio, A. P. to Lanxess Sybron Chemicals Inc. (2009). Process and apparatus for carbon dioxide capture via ion exchange resins. US8414689B2.
[32] Dow. (2019). Amberlite IR120 Technical data sheet. [online]. Available at: https://nshosting.dow.com/doc-archive/business/ier/ier_for_industrial_water_treatment/amberlite_ir120_h/tds/amberlite_ir120_h.pdf [Accessed 20 Nov. 2019].
[33] Dow. (2019). AMBERLITE® CG50 Product Data Sheet. [online] Available at: http://wetco-eg.com/wp-content/uploads/2017/12/AMBERLITE%C2%AE-CG50-Type-1-.pdf [Accessed 20 Nov. 2019].
[34] Fagerlund, G. (1973). Determination of specific surface by the BET method. Matériaux et Constructions, 6(3), pp.239-245.
[35] Malakhov, Y. and Samsonov, G. (1966). Regularities governing the thermal stability of oxides of the transition metals. *Soviet Powder Metallurgy and Metal Ceramics*, 5(12), pp.981-986.
[36] Fundamentals of Ion Exchange - DOW
[37] Xiao, B. and Thomas, K. (2005). Adsorption of Aqueous Metal Ions on Oxygen and Nitrogen Functionalized Nanoporous Activated Carbons. *Langmuir*, 21(9), pp.3892-3902.
[38] Rohm and Haas Amberlite IR-120 data sheet
[39] DOW Amberlite CG-50 data sheet
[40] Fundamentals of Ion Exchange – DOW
[41] Fundamentals of Ion Exchange – DOW
[42] Hawkes, S. (1996). All Positive Ions Give Acid Solutions in Water. *Journal of Chemical Education*, 73(6), p.516.
[43] Evan's pK_a table. Available at: http://evans.rc.fas.harvard.edu/pdf/evans_pKa_table.pdf
[44] Sienkiewicz, A., Krasucka, P., Charmas, B., Stefaniak, W. and Goworek, J. (2017). Swelling effects in cross-linked polymers by thermogravimetry. *Journal of Thermal Analysis and Calorimetry*, 130(1), pp.85-93.

[45] Socrates, G. (2001). *Infrared and Raman characteristic group frequencies*. Chichester, England: Wiley.pg 269
[46] Socrates, G. (2001). *Infrared and Raman characteristic group frequencies*. Chichester, England: Wiley. pg 221
[47] Socrates, G. (2001). *Infrared and Raman characteristic group frequencies*. Chichester, England: Wiley.pg 126
[48] Socrates, G. (2001). *Infrared and Raman characteristic group frequencies*. Chichester, England: Wiley.pg 127
[49] Socrates, G. (2001). *Infrared and Raman characteristic group frequencies*. Chichester, England: Wiley.pg 128
[50] Stahlbush, J. and Strom, R. (1990). A decomposition mechanism for cation exchange resins. *Reactive Polymers*, 13(3), pp.233-240.
[51] Teyssié, M. and Teyssié, P. (1961). High polymer chelates. I. Structure and stability of polymethacroylacetone chelates. *Journal of Polymer Science*, 50(154), pp.253-264.
[52] Quantochrome iQ3 gas sorption system operating manual

Dehydrogenation of Formic Acid with Au-Pd Nanoparticles on Titanate Nanotubes

Temi Majekodunmi and Yasmin Wulkau

Department of Chemical Engineering, Imperial College London, U.K.
December 2019

Abstract

The catalytic decomposition of formic acid is seen as a promising pathway for in situ production of H_2 in fuel cells. Au-Pd nanoparticles (NPs) on titanate nanotubes (Ti-NTs) made by sol-immobilisation have displayed high catalytic activities for selective oxidation. In this report, these Au-Pd NPs on Ti-NTs are investigated for the dehydrogenation of formic acid. After proving palladium (Pd) to be a more superior catalyst to various other metals, Pd/Ti-NTs was compared to Pd on SiO_2 and Al_2O_3 supports to evaluate the effect of the supports on catalytic activity. Au, Pd and Au-Pd on Ti-NTs were subsequently compared to determine whether the addition of Au to Pd improves the catalytic activity. The reaction proceeded in a vessel connected to a gas burette which was used to measure the volume of gas evolved over the reaction time. The amount of gas produced facilitated the calculation of formic acid conversion. GC analysis was carried out to determine if there was any CO production, and HPLC analysis was employed as a method to corroborate the formic acid conversion. It was discovered that there was a high degree of formic acid adsorption on the catalyst supports, which prevented high conversions from being achieved. Pd/Al_2O_3 gave an initial rate that was about 6 times greater than the other Pd catalysts which shows that Al_2O_3 was the superior support. Au-Pd on Ti-NTs exhibited a higher catalytic activity than Pd and Au on Ti-NTs which shows a synergy between Au and Pd. It was concluded that Au-Pd/Ti-NTs is an active catalyst for the dehydrogenation of formic acid. However, for use in fuel cell applications, its selectivity must be improved.

Keywords: Titanate nanotubes, Pd nanoparticles, Catalytic dehydrogenation, Formic acid, Hydrogen production

1 Introduction

As the alarming effects of global warming become more apparent, it is necessary to find greener and more sustainable sources of energy. With many countries committing to the shift towards renewable energy and away from fossil fuels, there is a requirement for the development of advanced energy storage technologies. Extensive research on renewable energy sources, including wind, solar and hydropower, has been conducted. However, such sources usually have geographic limitations and tend to present intermittent energy supply issues due to the variability of the sources.

Hydrogen is a source of energy that does not suffer from these intermittency challenges. As a result of being the most abundant element on earth, as well as having a high specific energy content, it presents a very attractive option as a fuel for widespread use. The hydrogen fuel cell is a promising technology, which uses the electrochemical reaction of hydrogen oxidation to form water and electricity. Unfortunately, the hydrogen fuel cell is subject to a number of limitations, the major ones being the handling, transportation and safe storage of hydrogen.

Hydrogen has a very low volumetric density so despite having a high energy content per unit mass ($120 MJkg^{-1}$), its energy density is relatively low (0.0118 MJ/L at room temperature). For this reason, hydrogen is typically stored as a highly compressed gas (5.32MJ/L) or liquid at temperatures as low as 20K ($10.0 MJ/L^1$). These forms of storage are energy intensive and pose both practical and safety concerns.[2] To address these concerns, alternative hydrogen storage by physisorption or chemical methods in various materials has been the subject of a considerable amount of research.

A particularly promising material as means of hydrogen storage is the simplest of the carboxylic acids, formic acid. Formic acid is a nontoxic compound with a high hydrogen density of 7.6MJ/L.[3] It is a liquid at room temperature which makes its transportation and use as a source of hydrogen more efficient, economical and safe.

Formic acid is known to decompose via a dehydrogenation reaction in the presence of a catalyst to form hydrogen and carbon dioxide in equimolar proportions.

$$HCOOH \rightleftharpoons H_2 + CO_2 \qquad \Delta G = -48.4\,kJ/mol$$

Selective catalysts must be used for this reaction under mild conditions and fairly low temperatures, to prevent the molecule from decomposing via the alternative dehydration pathway. This reaction forms carbon monoxide and water, as follows.

$$HCOOH \rightleftharpoons H_2O + CO \qquad \Delta G = -28.5\,kJ/mol$$

Carbon monoxide production must be avoided due to its high toxicity, alongside its ability to poison the catalysts used in fuel cells. Concentrations of carbon monoxide should be kept below 50ppmv.[4]

Another important reaction to take note of is the water-gas-shift reaction (WGSR). Due to the reversibility and coupling of the dehydration and dehydrogenation reactions, formic acid acts as an intermediate of the WGSR.[5,6]

$$H_2O + CO \rightleftharpoons HCOOH \rightleftharpoons H_2 + CO_2$$

This reaction has been exploited on an industrial-scale to obtain hydrogen gas, using selective catalysts in the presence of steam. At lower temperatures, the forward reaction is favoured, as the reaction equilibrium lies on the right. Also the addition of water in the system moves the equilibrium to the right and has been known to increase the dehydrogenation selectivity[7] .

This report focuses on the catalytic performance of metallic nanoparticles (NPs) for the selective dehydrogenation of formic acid.

2 Background

Metallic nanoparticles (NPs) are widely used in heterogeneous catalysis, and specifically in the dehydrogenation of formic acid.[8] Due to their small sizes, NPs have high surface area to volume ratios making them generally more catalytically active than their bulk metal counterparts.[9,10] Nanoparticles are typically supported on high surface area catalyst supports as these aids their dispersion and homogeneous distribution, increasing their exposure to substrate molecules in reaction mixtures.[11,12]

Palladium (Pd) is generally the most widely-used catalyst for the dehydrogenation of formic acid and dehydrogenation reactions in general.[12–15] However, due to their high reactivity, palladium monometallic catalysts often suffer from deactivation caused by the adsorption of reaction intermediates onto their surfaces.[11] Extensive research has been performed on the addition of less reactive metals, such as gold (Au) and silver (Ag), to Pd in order to combat this problem. The addition of such elements also increases the selectivity towards the dehydrogenation pathway and improve the catalytic activity.[8,11,13–15] Khawaji and Chadwick present results in their paper that show gold-palladium (Au-Pd) nanoparticles on titanate nanotubes (Ti-NTs) made by sol-immobilisation to be highly active catalysts for the oxidation of benzyl alcohol. This is due to the high dispersion of metal on the Ti-NT external surfaces and the narrow particle size distribution.[16] It is therefore expected that such catalysts are highly active for other reactions.

The main aim of this project was to determine whether these metals immobilised on Ti-NTs are active catalysts for the dehydrogenation of formic acid. The activity of Pd as a catalyst was first compared with various metals that have been widely studied for this reaction, namely platinum (Pt), ruthenium (Ru), and Iridium (Ir)[12–15,17,18] in order to verify the superiority of Pd. The activity of Pd on Ti-NTs was then compared with Pd on silica (SiO_2) and Pd on alumina (Al_2O_3). Pd on both SiO_2 and Al_2O_3 supports have been shown to catalyse the reaction with high activities and selectivities towards the dehydrogenation pathway.[19] This comparison allows for the evaluation of the effectiveness of Ti-NTs as catalyst supports for this particular reaction.

Finally, the catalytic activity of Pd/Ti-NTs, Au/Ti-NTs and Au-Pd/Ti-NTs were studied to determine if the addition of Au to Pd increases its catalytic activity and selectivity, or if antisynergy exists between the two metals. While there seems to be a general consensus that the addition of less reactive metals like Ag or Au to Pd improves catalytic activity,[8,11,14,15] Carter et al claim that there is strong antisynergy between Au and Pd for the dehydrogenation of formic acid.[20] This analysis will verify one of these assertions.

3 Method

Chemicals

A >98 wt. % solution of formic acid was purchased from Sigma-Aldrich and diluted with deionised water to the required concentration. A 5 wt. % Pd on Al_2O_3 catalyst powder was purchased from Sigma-Aldrich. The other catalysts used, 1.5% Pd on Ti-NTs, 1% Au on Ti-NTs, 0.65% Au 0.35% Pd on Ti-NTs, 5% Pd on SiO_2, 5% Pt on SiO_2, 5% Ru on SiO_2 and 5% Ir on SiO_2, were prepared and synthesised in the laboratory.

Reactor Set-up

A schematic of the reactor set-up is displayed in Figure 1. A Büchi AG 50mL reaction vessel, able to withstand pressures of up to 8 bar, was connected to an inlet, an outlet, and a septum for injection and gas sampling. The reactor was immersed in a water bath; this sat on a hot plate which controlled the temperature at which the reactor was operating at. The hot plate (IKA C-MAG HS10 Digital) was magnetic, enabling the use of a magnetic stirrer in the reactor. It was necessary to calibrate the temperatures between the water bath and the inside of the reactor, as the thermocouple could not be used during the reaction due to a risk of corrosion by formic acid.The calibration data is shown in the supplementary information (Figure 11).

The gas burette was initially filled with DI water and submerged into a beaker. The burette was used to measure the gas evolved from the decomposition, fed by a flexible tube attached to the reactor outlet. Temperature indicators allowed for the temperatures of the water

bath and the DI water in the beaker to be monitored and recorded.

Nitrogen was used to purge the system before the reaction and to dilute the H_2 exit stream following the termination of the reaction. Valves controlled the flow of nitrogen gas into and out of the reactor system when required.

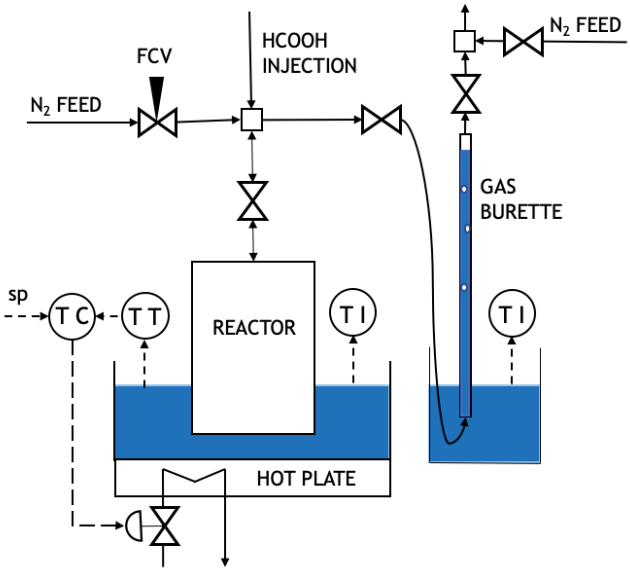


Figure 1: Schematic of Reactor Set-up

Reaction Procedure

Over time, catalysts are prone to oxidation so before use, all the catalysts were reduced by heating in a gas stream of 5 wt. % hydrogen in nitrogen at 250°C for two hours.

For each run, 9mL of deionised water and a measured mass of catalyst was added to the reaction vessel and magnetically stirred with a stirrer setting of 3, which equates to approximately 800 RPM. The setup was left for 30 minutes in order to allow the water bath and reactor contents to reach a steady temperature. During this time, the reaction vessel was flushed with nitrogen five times to remove any air from the system. 1mL of a 1M solution of formic acid was injected into the reaction vessel, which marked the start of the reaction. Gas evolution was recorded over time, as well as the temperature of the water bath that housed the reaction vessel, and the temperature of the water being displaced by the product gases. The reaction was monitored over 2 hours. The amount of gas produced facilitated the calculation of formic acid conversion in this time.

Chromatography Analysis

Following the termination of the reaction, a 250μL sample of the reaction gas was obtained and analysed by gas chromatography (GC), using helium as the carrier gas. The GC-2014 Shimadzu model with a flame ionisation detection unit was used. This was fitted with a Agilent Technologies, Inc column (30m x 0.320mm ID with 20 micron pores). GC analysis was used to determine whether any carbon monoxide was being produced alongside carbon dioxide. Calibration curves were obtained using known gas standards (Supplementary information, Figures 8 and 9).

High Pressure Liquid Chromatography (HPLC) was used to corroborate the conversion of formic acid with that obtained from gas burette readings. The reaction liquid was collected from the reactor and placed into a centrifuge bottle. The mixture was centrifuged for 45 minutes at 6500 RPM, in order to separate the catalyst from the unconverted formic acid and DI water. A syringe filter removed any remaining solid particles which would block the chromatography column. The CTO-20AC, Shimadzu, Japan model used was fitted with a SUPELCOGEL$_{TM}$ C-610H HPLC column (30cm x 7.8mm with 0.45 micron pores and a 25mm filter). 0.1 vol. % phosphoric acid was employed as the mobile phase. A calibration curve was obtained using formic acid standards (Supplementary information, Figure 10). 50μL of citric acid was used as an internal standard and added to the 1mL liquid samples which were analysed.

Calculations & Corrections

In order to calculate the conversion of formic acid using gas burette readings, a few corrections were applied and assumptions made. The gas produced from the reaction was assumed to be ideal; this was considered to be reasonable given the relatively low pressure (1 atm) the gas was being produced and collected at.

It is known that carbon dioxide is soluble in water to form carbonic acid.

$$H_2O + CO_2 \longrightarrow H_2CO_3$$

The acid then dissociates as follows:

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3{}^- \rightleftharpoons H^+ + CO_3{}^{2-}$$

The solubility of carbon dioxide into water would yield an underestimation for the extent of formic acid conversion. As a means to determine the amount of carbon dioxide that dissolved in the water, the pH of the water was recorded before and after multiple reaction runs. The amount of carbon dioxide dissolved was consistently found to be in the order of 10^{-8} mol. As such, the effect of carbon dioxide solubility was deemed to have a negligible effect on conversion calculations.

Furthermore, it was observed that the level of the water dropped in the gas burette, even in the absence of a reaction. The exact reasoning for this is unknown, but potential reasons may be due to gas leakage or the gas components dissolving into the liquid. Multiple runs were carried out to account for this effect. A gas burette was filled with water and left for 12 hours, and the volume of gas that had displaced the water was noted.

A 'dropping water level' rate of 0.006 mL/min was calculated and this correction was applied to subsequent calculations.

GC showed that the volume of CO was essentially zero in comparison to the volume of carbon dioxide produced. As a result, the gas collected in the gas burette was assumed to be an equimolar mixture of carbon dioxide and hydrogen.

The following calculation was then used to determine the formic acid conversion:

$$Conversion(\%) = \frac{\text{Moles of } H_2 \text{ produced}}{\text{Initial moles of formic acid}} \times 100\% \quad (1)$$

$$n(H_2) = \frac{Volume(H_2)(m^3) \times Pressure(Pa)}{8.3145 JK^{-1}mol^{-1} \times Temperature(K)} \quad (2)$$

The HPLC calibration curve was used to determine the mass ratio of formic acid in the reaction liquid from the peak areas on the chromatogram. This allowed the conversion of the formic acid to be calculated as follows:

$$Conversion(\%) = \frac{n(HCOOH)_0 - n(HCOOH)}{n(HCOOH)_0} \times 100\% \quad (3)$$

n(HCOOH) = moles of formic acid remaining in solution and $n(HCOOH)_0$ = initial moles of formic acid in solution = 10^{-3} mol.

Repeats of the experiments were carried out, and the standard deviation of the data was calculated. The standard deviation was used as the magnitude of the uncertainty in the results.

4 Results

Mass Transfer Limitations

To ensure there were no mass transfer limitations, different masses of Pd on Al_2O_3 support were tested (20 mg, 30 mg and 40 mg). Pd on Al_2O_3 was available in abundance so multiple runs were possible. A graph was plotted of gas evolved against time. It was assumed that the rate of reaction was first order; from this, an initial rate was determined. As shown in Figure 2, there were no mass transfer limitations in this mass range at the chosen stirrer speed. Consequently, only masses in this range were tested for subsequent experimental runs.

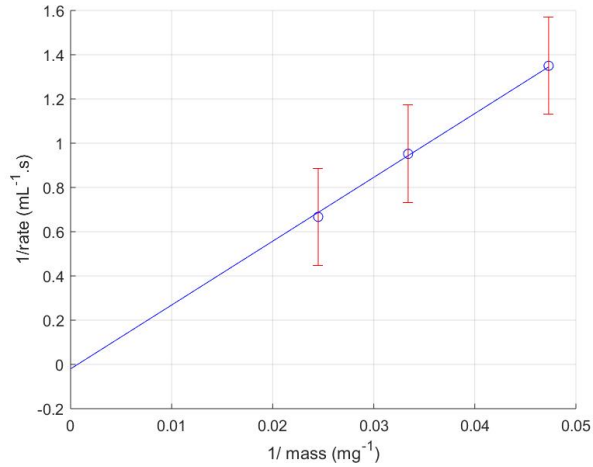


Figure 2: Determination of mass transfer limitations for mass range

Effect of Temperature

In order to test the temperature-dependence of the reaction, 21mg of Pd on Al_2O_3 was tested at various temperatures. The results of this can be shown in Figure 3. As expected, a larger initial rate was achieved at higher temperatures. For all three temperatures tested, a conversion of 22% was achieved. As essentially the same conversion was achieved for all temperatures tested, a water bath temperature of 40°C was chosen arbitrarily for all subsequent reactions.

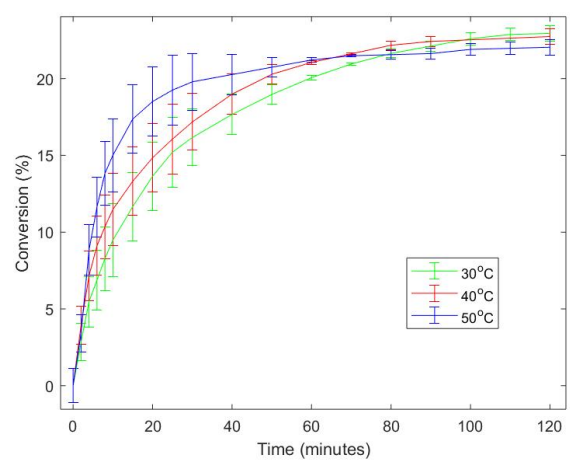


Figure 3: Determination of mass transfer limitations for mass range

Palladium Catalyst

Figure 4 shows that Pd yielded the highest initial rate of all the metal catalysts on SiO_2 tested at 40°C. In fact, the initial rates for all the other catalysts tested were very low even though some conversion of formic acid eventually took place. This was verification that Pd was the most active catalyst, as proven by many others.[11,13]

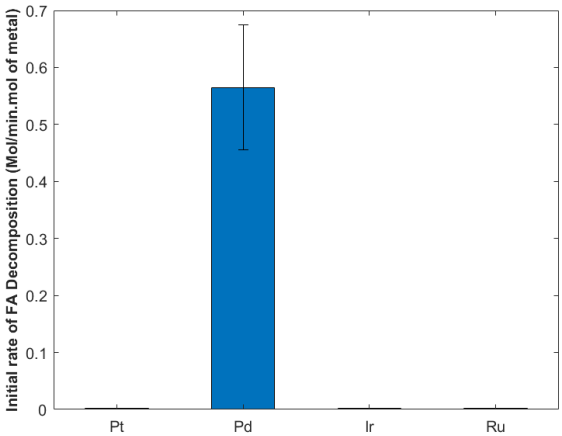


Figure 4: Initial rate bar chart for metal catalysts on SiO_2 at 40°C

Catalytic Activity and Efficiency

Initial Rates

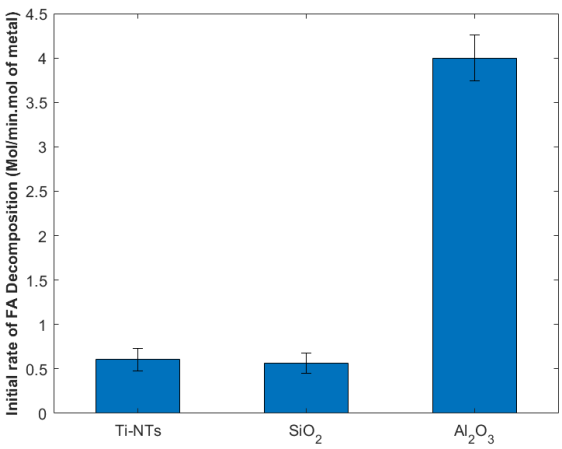


Figure 5: Initial rate bar chart for Pd catalysts

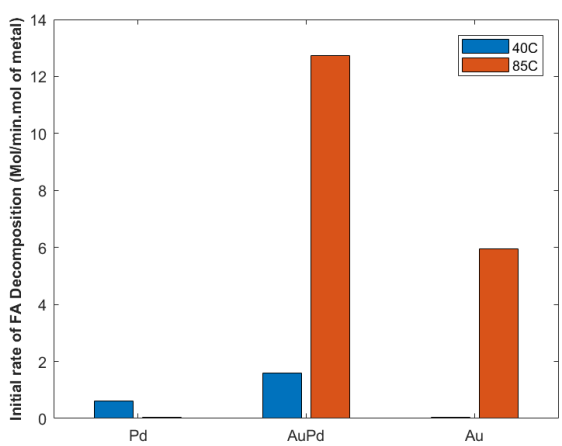


Figure 6: Initial rate bar chart for catalysts on Ti-NTs

The initial rates per mole of active metal was calculated for each run and the results are shown in Figures 5 and 6. Initial rate here is the moles of hydrogen produced per unit time in the first 4 minutes of the reaction.

Conversions

The conversion of formic acid was first calculated using the values for the volume of gas produced from the reaction. HPLC analysis was used as a means to corroborate these conversions. Results are shown in Table 1.

Catalyst	T (°C)	% Conversion	
		Burette	HPLC
Pd on Al	40	22.5 ± 1.52	44.7
Pd on SiO_2	40	5.18 ± 0.36	32.7
Au on Ti-NTs	40	0	11.49
Au-Pd on Ti-NTs	40	5.37 ± 2.63	29.5
Pd on Ti-NTs	40	3.68 ± 0.73	20.6
Au on Ti-NTs	85	1.62 ± 1.13	26.4
Au-Pd on Ti-NTs	85	7.50 ± 0.53	33.1
Pd on Ti-NTs	85	0	8.82

Table 1: Conversion of HCOOH

Turnover Number

The turnover number (TON) has been defined as follows.

$$TON = \frac{\text{Mol of } HCOOH \text{ converted}}{\text{Mol of catalyst used}} \quad (4)$$

Catalyst	T (°C)	TON
Pd on Al	40	22.5±1.5
Pd on SiO_2	40	9.5±0.5
Au on Ti-NTs	40	0
Au-Pd on Ti-NTs	40	37.9±26.2
Pd on Ti-NTs	40	21.5±9.4
Au on Ti-NTs	85	11.9±10.6
Au-Pd on Ti-NTs	85	53.5±3.7
Pd on Ti-NTs	85	0

Table 2: TON of various catalysts

CO_2 Selectivity

The relative selectivity of carbon dioxide to carbon monoxide, $R_{CO2/CO}$ produced can be calculated using results from the GC analysis, and the following equation.

$$R_{CO2/CO} = \frac{\text{Mol of } CO_2 \text{ produced}}{\text{Mol of CO produced}} \quad (5)$$

Catalyst	T (°C)	[CO] (ppmv)	$R_{CO2/CO}$
Pd/Al_2O_3	40	23.5	21,300
Pd/SiO_2	40	151	3,310
Au/Ti-NTs	40	-	-
Au-Pd/Ti-NTs	40	424	1180
Pd/Ti-NTs	40	83,300	6
Au/Ti-NTs	85	1,840	272
Au-Pd/Ti-NTs	85	608	822
Pd/Ti-NTs	85	-	-

Table 3: Relative selectivity of various catalysts towards the dehydrogenation pathway, where T is temperature

5 Discussion

Carbon Imbalance

From the results shown in Table 1, there was a clear difference in the conversion values. HPLC conversions were consistently higher than conversions calculated from the volume of hydrogen production. From this, the rate of formic acid consumption did not appear to equal the rate of hydrogen production, violating mass conservation.

Two factors were considered in order to explain this difference: either formic acid was adsorbing onto the catalyst support without reacting, or formic acid was reacting via dimerisation or polymerisation to form another product. In the HPLC used, size-exclusion was the method for component separation. If there was any dimerisation or polymerisation, peaks would have appeared at a lower retention time than the formic acid peak. However, no other peaks appeared on the HPLC chromatogram aside from the formic acid peak and the internal standard peak. As such, it was concluded that no dimerisation or polymerisation occurred.

To verify whether adsorption was occurring, the formic acid was added to catalyst supports in the absence of metal and was subjected to the same experimental conditions as the metal catalysts were. This was done at 40°C, and the resulting reaction mixture was analysed by HPLC. Although no gas evolution occurred, the concentration of formic acid from this analysis was found to be less than initial concentration of the reaction liquid at t=0. The results (Table 4) show that at 40°C, there is significant adsorption on all the catalyst supports. It should be noted that the extent of adsorption on each support differed.

Support	% HCOOH adsorbed
SiO_2	44.0
Al_2O_3	28.8
Ti-NTs	12.5

Table 4: Adsorption on various supports

It is expected that the rate of desorption increases at higher temperature. This was demonstrated qualitatively with a run of Au-Pd catalyst at 40°C. When the experiment ran at 40°C, 0.7mL of gas was produced in 30 minutes but when the temperature was increased to 50°C 2.3mL of gas was produced 2.3mL in 30 minutes.

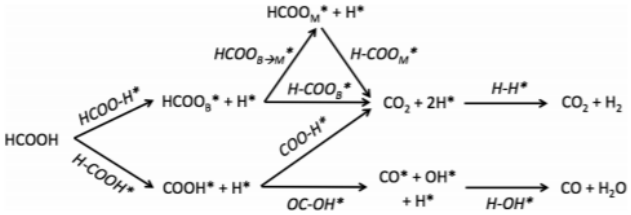
Point of Zero Charge (PZC) corresponds to the pH when the net charge of the total catalyst surface is zero. To gain more insight into the method of adsorption that was occurring, the PZC values of the catalysts' supports were compared with the pH of the reaction liquid (which varied from pH 6.71 to pH 7.02). The PZC values of the support materials are as follows: 5.5 for Ti-NT,[16] 2.2 for SiO_2,[21] and 9.1 for Al_2O_3.[22] The pH exceeded the PZC values on the Ti-NT and SiO_2 surfaces, implying that the catalyst surfaces were negatively charged. As a result, electrostatic adsorption can be ruled out, given that formic acid dissociates into negatively-charged formate ions. The PZC value of Al_2O_3 is higher than the pH at the reaction conditions. This would suggest that the Al_2O_3 was positively-charged, but due to formic acid being a weak acid, the extent of dissociation would be extremely small. For both these reasons, chemisorption provides the most reasonable mechanism of adsorption.

In summary, formic acid adsorption onto the catalyst support appears to be the reason for the discrepancy in conversion results. To fully understand the effects of adsorption on this reaction, further analysis would have to be carried out with the catalysts. The effect of temperature on the rate of adsorption and desorption should be studied in greater detail.

Effects of Supports on Catalytic Performance

To determine the effect of the supports on the catalytic performance, the activities of the Pd catalysts (Pd/Ti-NTs, Pd/SiO_2 and Pd/Al_2O_3) were compared. The initial rate of Pd/Al_2O_3 support was 4 mol.min^{-1}.(mol of metal)$^{-1}$, approximately six times greater than initial rates for Pd/SiO_2 and Pd/Ti-NTs, as shown in Figure 5. The formic acid conversion obtained using Pd/Al_2O_3 was also greater than that using Pd/SiO_2 and Pd/Ti-NTs. Both of these metrics would indicate that Al_2O_3 is the superior catalyst support in this analysis.

A potential explanation for this is the reactivity of the Al_2O_3 support. Al_2O_3 has been shown to react with formic acid at lower temperatures to produce formate ions,[23,24] which is an intermediate in the dehydrogenation of formic acid.[25,26] The decomposition of formic acid is initiated by activating either the O-H or C-H bond of the HCOOH to form formate (HCOO*) or carboxyl (COOH*) species on the catalyst surface. Figure 7 shows that the formate pathway only leads to dehydrogenation while the formation of the carboxyl species could lead to dehydration. Therefore, if the presence of Al_2O_3 in solution increased the production rate of formate ions, this would explain the higher initial rate of reaction.



[a]Reaction species shown in *italic font* are the transition states, and those marked with an asterisk (*) are the reaction intermediates adsorbed on the surface.

Figure 7: Formic acid decomposition pathway[25]

Another reason for Al_2O_3 exhibiting higher initial rates could be due to Pd leaching into the reaction mixture. It has been noted that Pd has a tendency to leach into solutions. For this reaction, leaching would result in formic acid decomposition occurring in the solution by means of homogeneous catalysis, as opposed to heterogeneous catalysis, leading to an increase in the catalytic activity. Leaching was not studied within this project but it is possible that Al_2O_3 releases Pd NPs into solution as Pd/Al_2O_3 has been shown to display this effect.[27] Microwave Plasma analysis could be used to analyse the reaction liquid to verify this.[15]

Table 4 shows that SiO_2 adsorbed the most formic acid of the three supports. This is likely to be the reason that Pd/SiO_2 gave the lowest conversion, initial rate and TON of the Pd catalysts.

Pd on Ti-NTs and Al_2O_3 have similar TON values, both being significantly higher than that of SiO_2 at 40°C. TON is a measure of the catalytic efficiency. Although the initial rate of Pd/Ti-NTs was not as high as Pd/Al_2O_3, the reason the TONs are essentially equivalent may be due to the metal particles on Ti-NTs being more highly dispersed than on the Al_2O_3 support.[16] It would be insightful to characterise the catalyst surfaces using x-ray diffraction (XRD) to examine the differences in metal particle size and dispersion of NPs on the supports. It is expected that smaller particles promote the rate of reaction due to a higher surface area to volume ratio, but this may not be the case. An XRD analysis would provide a clearer understanding regarding the differences in catalytic activities.

Effects of Au on the Activity of Pd Catalysts

To determine if there was antisynergy between Au and Pd, the activities of the catalysts on Ti-NTs supports were compared. These were Au, Pd and Au-Pd. They were reacted at two temperatures: 40°C to be able to compare with the other runs in this project and 85°C, the temperature Carter et al performed their experiments at.[20]

From Table 1 and Figure 6, it can be shown that synergy exists between Au and Pd. This is because the initial rate of reaction, conversion, and TON are all higher for Au-Pd in comparison to Au and Pd at both temperatures tested. This directly contradicts the results of Carter et al,[20] but agrees with other findings.[8,11,15]

Unexpectedly, Pd/Ti-NTs was inactive at 85°C. There was no observable reaction for any of the runs that were performed. This could be due to extensive adsorption of formic acid on the catalyst support. To verify this, the catalyst separated from the reaction liquid after a run at 85°C was placed in the reactor setup with 9mL of DI water. The temperature was increased in steps every 30 minutes from 50°C up until 90°C. Table 5 shows the amount of gas that was produced at each temperature.

T (°C)	Gas evolved (mL)
50	1.5
60	2.6
70	3.7
85	6
90	2.5

Table 5: Gas evolved at different temperatures using recycled Pd/Ti-NTs

The data from the table shows that there was a significant amount of formic acid adsorbed on the catalyst, as the desorbed formic acid from the support was able to decompose on the catalyst active sites. A total of 16.4mL of gas was evolved; this corresponds to a formic acid conversion of 34.1%. This does not tally with the HPLC data for Pd/Ti-NTs at 85°C (which gives a conversion value of 8.82% (Table 1)) and adsorption data (which shows that Ti-NTs adsorb the least formic acid (Table 4)). It is expected that the effect of adsorption would be lower at higher temperatures, so these results are unusual. It should be noted that the Ti-NTs tested in Table 4 were not identical to the supports used in the Pd/TiNTs catalyst. Further analysis would have to be carried out on Pd/Ti-NTs at 85°C to determine why it was inactive and why there was so much adsorption of formic acid on the supports.

The results above demonstrate the synergy between Au and Pd on Ti-NTs. However, for a clearer insight into the interaction between the two elements, intermediate Au-Pd compositions would have to be tested.

Relative Selectivity Towards Dehydrogenation Reaction

All catalysts tested favoured the dehydrogenation pathway over the dehydration pathway, as shown by the large amount of carbon dioxide in comparison to carbon monoxide produced in the reaction. The values calculated for the relative selectivity are shown in Table 3.

The $R_{CO2/CO}$ value obtained for Au-Pd/Ti-NTs support is significantly better than that of Pd/Ti-NTs support. This agrees with findings in several papers stating that the addition of Au to Pd improves the catalyst selectivity towards the dehydrogenation pathway.[8,15]

Pd/Al_2O_3 supports have a higher selectivity towards the dehydrogenation pathway than Pd on either SiO_2 or Ti-NTs at 40°C. This could again be explained by the reaction of Al_2O_3 with formic acid to produce formate ions.

As determined in several studies, concentrations of carbon monoxide should be kept below 50ppmv to prevent the poisoning of the catalytic surface by carbon monoxide.[4] For Pd/Al_2O_3 supports, the concentration of carbon monoxide in the reactor was below 50ppmv; this degree of CO production would be acceptable in fuel cells. The concentration of CO produced in reactions using Pd/Si_2O_3 and catalysts on Ti-NTs exceeded the 50ppmv threshold. Pd on Ti-NT showed a substantially lower selectivity than all other catalysts tested. However, this does agree with findings that Au improves the catalyst selectivity. Consequently, Au-Pd catalysts with higher gold contents than tested in this project may prove to be more suitable in fuel cell applications.

TON as a Catalytic Activity Metric

In this paper, TON was used as a measure of catalytic efficiency. However, in heterogeneous catalysis, turnover frequency (TOF) is a widely-used metric to compare the efficiencies of different catalysts. TOF is calculated using the following equation.

$$TOF(s^{-1}) = \frac{\text{Molecules of HCOOH converted}}{\text{Number of active sites} \times \text{reaction time (s)}}$$

Brunauer-Emmett-Teller (BET) analysis could be performed to determine the number of catalyst active sites, and therefore TOF. The results obtained from this could be compared to literature values.

6 Conclusions

A study of initial rates proved Pd to be the most superior catalyst of all the metal catalysts on SiO_2, having an initial rate of formic acid decomposition of 0.6 mol.min^{-1}.(mol of catalyst)$^{-1}$. Platinum, iridium, and ruthenium all yielded initial rates of 0 mol.min^{-1}.(mol of catalyst)$^{-1}$ at 40°C.

The Al_2O_3 support exhibited the highest catalytic activity for Pd in comparison to SiO_2 and Ti-NTs. The initial rate of Pd on Al_2O_3 support was 4 mol.min^{-1}.(mol of catalyst)$^{-1}$; this was approximately eight times higher than the initial rates on the other two supports. It is suspected that the leaching of Pd into solution and the reactivity of Al_2O_3 with formic acid may have accounted for the higher catalytic activity and selectivity.

It was also discovered that the addition of Au to Pd resulted in an enhanced catalytic activity. This implies the existence of a synergy between Au and Pd. The bimetallic Au-Pd catalyst tested comprised 65% Au and 35% Pd. It would be beneficial to test a wider variety of proportions for this bimetallic catalyst to allow for the optimum metal proportions to be identified.

In all experimental runs, adsorption of formic acid onto the catalyst support prevented higher conversions of formic acid from being achieved. It should be noted that the effect of adsorption on conversion is more of a problem in batch processes such as this one. In continuous flow reactors, adsorption would be less of an issue as the catalyst supports would become saturated with formic acid and reaction would be able to proceed on the catalyst's active sites.

It would be useful to conduct catalyst recyclability tests. This would allow for the evaluation of the reusability of the catalysts, which is an essential consideration in industrial processes. It is expected that the activity will decrease but the extent of this is unknown. Potential reasons for the reduction in catalytic activity are as follows: leaching of Au and Pd nanoparticles into the solution, CO poisoning of the active sites, and adsorption of formic acid onto the supports. In order to quantitatively assess the reduction in catalytic activity, the catalyst would be recovered from the reaction liquid by filtration. The filter cake would be dried, enabling the catalyst to be reused. The activity of the reused catalyst can then be compared to the activity of the fresh catalyst.

In conclusion, it was determined that Au-Pd/Ti-NTs is an active catalyst for the dehydrogenation of formic acid. However, in order for this catalyst to be used in fuel cell applications, its selectivity has to be improved to reduce the amount of CO produced.

7 Acknowledgements

We would like to thank Yiming Zhang and Senpei Peng for the help they provided to us during this project.

References

[1] College of the Desert. Hydrogen fuel cell engines and related technologies, 2001.

[2] Saba Niaz, Taniya Manzoor, and Altaf Hussain Pandith. Hydrogen storage: Materials, methods and perspectives. *Renewable and Sustainable Energy Reviews*, 50:457 – 469, 2015.

[3] DeepResource. Formic acid as car fuel.

[4] Xuan Cheng, Zheng Shi, Nancy Glass, Lu Zhang, Jiujun Zhang, Datong Song, Zhong-Sheng Liu, Haijiang Wang, and Jun Shen. A review of pem hydrogen fuel cell contamination: Impacts, mechanisms, and mitigation. *Journal of Power Sources*, 165(2):739 – 756, 2007. IBA – HBC 2006.

[5] Ken Yoshida, Chihiro Wakai, Nobuyuki Matubayasi, and Masaru Nakahara. Nmr spectroscopic evidence for an intermediate of formic acid in the watergasshift reaction. *The Journal of Physical Chemistry A*, 108(37):7479–7482, 2004.

[6] Yoshiro Yasaka, Ken Yoshida, Chihiro Wakai, Nobuyuki Matubayasi, and Masaru Nakahara. Kinetic and equilibrium study on formic acid decomposition in relation to the water-gas-shift reaction. *The Journal of Physical Chemistry A*, 110(38):11082–11090, 2006. PMID: 16986841.

[7] Production of co-free h2 from formic acid. a comparative study of the catalytic behavior of pt metals on a carbon support. *Journal of Catalysis*, 279(1):213 – 219, 2011.

[8] Dehydrogenation of formic acid by heterogeneous catalysts. *CHIMIA International Journal for Chemistry*, 69(6), 2015.

[9] Wei Xie and Sebastian Schlücker. Surface-enhanced raman spectroscopic detection of molecular chemo- and plasmo-catalysis on noble metal nanoparticles. *Chem. Commun.*, 54:2326–2336, 2018.

[10] Ahmad Alshammari, V. Narayana Kalevaru, and Andreas Martin. Metal nanoparticles as emerging green catalysts. *Green Nanotechnology - Overview and Further Prospects*, 2016.

[11] Miriam Navlani-García, Kohsuke Mori, David Salinas-Torres, Yasutaka Kuwahara, and Hiromi Yamashita. New approaches toward the hydrogen production from formic acid dehydrogenation over pd-based heterogeneous catalysts. *Frontiers in Materials*, 6:44, 2019.

[12] R. S. Oosthuizen and V. O. Nyamori. Carbon nanotubes as supports for palladium and bimetallic catalysts for use in hydrogenation reactions. *Platinum Metals Review*, 55(3):154–169, Jan 2011.

[13] Karaked Tedsree, Tong Li, Simon Jones, Chun Wong Aaron Chan, Kai Man Kerry Yu, Paul A. J. Bagot, Emmanuelle A. Marquis, George D. W. Smith, and Shik Chi Edman Tsang. Hydrogen production from formic acid decomposition at room temperature using a ag–pd core–shell nanocatalyst. *Nature Nanotechnology*, 6(5):302–307, Oct 2011.

[14] Simon Jones, Amy Kolpin, and Shik Chi Edman Tsang. Modification of pd for formic acid decomposition by support grafted functional groups. *Catalysis, Structure & Reactivity*, 1(1):19–24, 2015.

[15] Felipe Sanchez, Davide Motta, Alberto Roldan, Ceri Hammond, Alberto Villa, and Nikolaos Dimitratos. Hydrogen generation from additive-free formic acid decomposition under mild conditions by pd/c: Experimental and dft studies. *Topics in Catalysis*, 61(3):254–266, Apr 2018.

[16] Motaz Khawaji and David Chadwick. Au-pd bimetallic nanoparticles immobilised on titanate nanotubes: A highly active catalyst for selective oxidation. *ChemCatChem*, 9(23):4353–4363, 2017.

[17] Katerina Sordakis, Conghui Tang, Lydia K. Vogt, Henrik Junge, Paul J. Dyson, Matthias Beller, and Gábor Laurenczy. Homogeneous catalysis for sustainable hydrogen storage in formic acid and alcohols. *Chemical Reviews*, 118(2):372–433, 2018. PMID: 28985048.

[18] Céline Fellay, Ning Yan, Paul J. Dyson, and Gábor Laurenczy. Selective formic acid decomposition for high-pressure hydrogen generation: A mechanistic study. *Chemistry – A European Journal*, 15(15):3752–3760, 2009.

[19] Lijun Jia, Dmitri A. Bulushev, and Julian R.H. Ross. Formic acid decomposition over palladium based catalysts doped by potassium carbonate. *Catalysis Today*, 259:453 – 459, 2016. Catalytic Science and Technology in Sustainable Energy.

[20] James H. Carter, Sultan Althahban, Ewa Nowicka, Simon J. Freakley, David J. Morgan, Parag M. Shah, Stanislaw Golunski, Christopher J. Kiely, and Graham J. Hutchings. Synergy and anti-synergy between palladium and gold in nanoparticles dispersed on a reducible support. *ACS Catalysis*, 6(10):6623–6633, 2016. PMID: 27990317.

[21] José Antonio Alves Júnior and João Baptista Baldo. The behavior of zeta potential of silica suspensions. *New Journal of Glass and Ceramics*, 04(02):29–37, 2014.

[22] Bimal P. Singh, Ruben Menchavez, Chika Takai, Masayoshi Fuji, and Minoru Takahashi. Stability of dispersions of colloidal alumina particles in aqueous suspensions. *Journal of Colloid and Interface Science*, 291(1):181 – 186, 2005.

[23] Ling-Yan Wu, Sheng-Rui Tong, Si-Qi Hou, and Mao-Fa Ge. Influence of temperature on the heterogeneous reaction of formic acid on -al2o3. *The Journal of Physical Chemistry A*, 116(42):10390–10396, 2012. PMID: 23025659.

[24] S. R. Tong, L. Y. Wu, M. F. Ge, W. G. Wang, and Z. F. Pu. Heterogeneous chemistry of monocarboxylic acids on alpha;-al_2o_3 at different relative humidities. *Atmospheric Chemistry and Physics*, 10(16):7561–7574, 2010.

[25] Jong Suk Yoo, Frank Abild-Pedersen, Jens K. Nørskov, and Felix Studt. Theoretical analysis of transition-metal catalysts for formic acid decomposition. *ACS Catalysis*, 4(4):1226–1233, 2014.

[26] Matthew D. Marcinkowski, Jilei Liu, Colin J. Murphy, Melissa L. Liriano, Natalie A. Wasio, Felicia R. Lucci, Maria Flytzani-Stephanopoulos, and E. Charles H. Sykes. Selective formic acid dehydrogenation on pt-cu single-atom alloys. *ACS Catalysis*, 7(1):413–420, 2017.

[27] Saeeda S. Soomro, Farzana L. Ansari, Konstantinos Chatziapostolou, and Klaus Köhler. Palladium leaching dependent on reaction parameters in suzuki–miyaura coupling reactions catalyzed by palladium supported on alumina under mild reaction conditions. *Journal of Catalysis*, 273(2):138 – 146, 2010.

A method to calculate interface properties for atomic force microscopy-imaged sessile droplets at a liquid-solid interface

George-Claudiu Savulescu
Department of Chemical Engineering, Imperial College London, U.K.

Abstract. Roughness is a key factor influencing wettability properties. Micro computed tomography (micro-CT) is commonly used to scan in porous media which have a large internal surface. However, this approach does not have sufficient resolution to account for roughness. Preliminary studies have demonstrated that AFM could be utilized as an alternative technique to characterise the wetting of the internal surface of porous media. The study includes the image analysis, covering volume, 2-phase surface area, 3-phase contact lines and contact angles of a model system (oil-brine-calcite), which manifests in form of water-droplets along the solid surface. Thereby different parameters in the image analysis, including filtering parameters and resolution were varied and compared to an idealized droplet. The error for the calculation of the volume, 2-phase interfaces and contact angles is low, while the 3-phase line has a high error. The contact angles have a wide quasisymmetric distribution in an interval between 5° and 35°, which reveals the degree of roughness on the contact line. The results show that AFM provides an accurate result for volume, surface area and contact angle, while contact line length remains resolution depended. This indicates that AFM is suitable to complement micro-computed tomography approaches for wetting determination through higher resolution on subsampled surfaces.

1.Introduction

1.1 Relation between roughness and wetting

Roughness represents an essential parameter for characterizing wetting, a property with significant practical relevance. For example, surface nanostructure impacts the wettability of jackets. Very small pores and a compact structure does not allow water molecules to impregnate in the solid surface and make the jacket water-repellent. On the other hand, sufficiently large pores are impregnated by water droplets when it rains and the jacket becomes wet. Structure of the surface thus influences the level of wettability of a jacket. Most solids are naturally rough, but coating can also cause roughness, specifically when the coating film produces microdroplets at the surface. Conversely, very few solids are molecularly flat. [1]. The contact angle is fixed univocally by the chemical nature of the different phases, but this can be severely affected if the solid is rough. Gibbs pointed out that defects on a solid, such as roughness or chemical heterogeneity can pin a contact line [1].Calcite is a chemically uniform surface and in our case, chemical heterogeneity does not apply. Due to the development of advanced techniques for analysing well-controlled nanotextures on a solid, the effects of roughness on wettability are being exploited to induce wetting properties, such as spontaneous filmification, superhydrophobicity, superoleophobicity, and interfacial slip, that could not be achieved without roughness. [1].The results of this study aim to characterize roughness of the solid surface, by calculating interface properties, which can be further used for the analysis of wettability. The main question that this work answers is: how sensitive is the measurement of fluid along solids with AFM and can AFM image provide more accurate wetting measures than micro-CT?

1.2.Background

1.2.1 Comparison between micro computed tomography (micro-CT) and AFM

Images have previously been processed with micro-CT. However, for rocks such as calcite, whose roughness can only be observed at a nanometric scale, micro-CT cannot identify roughness accurately. [2]. Micro-CT technique scans at a larger scale and therefore, water droplets at an oil-calcite interface are identified as droplets residing on a flat surface (Figure 1), rather than droplets residing on a rough surface, as it occurs in the real droplet (Figure 2). As a result, the contact angle distribution will have a high degree of error for this kind of measurement and will not identify the

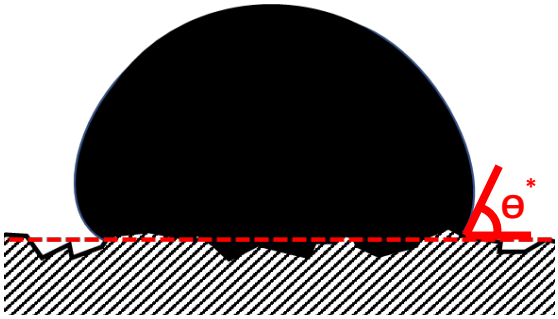


Figure 1: Sessile droplet on a solid surface with a degree of roughness- the red line illustrates the fact that micro-CT measures the rough nano-scale surface as flat

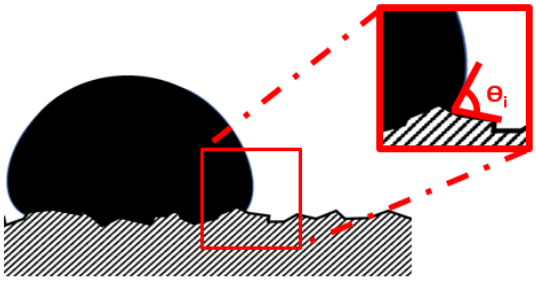


Figure 2: An AFM-imaged droplet includes the level of roughness of the nano-scale surface

properties of the interfaces which we study in this paper. Consequently, another technique was preferred for the scanning of the sample droplet. Atomic force microscopy scans at a significantly lower, nanometric scale, at which roughness of calcite can be accurately characterized. An example of a rough surface is represented in figure 2. In this case, contact angles can be measured accurately at each point along the contact line of the droplet, because AFM provides information about roughness at atomic level, to an order of nanometres.

1.2.2. Characterization of the sample

This study starts from the image in Figure 3.a., where an AFM-scanned adhesion map of the sample surface used for this study is represented. The brown surface represents the calcite, the yellow surfaces are the droplets and the surface above calcite and the droplets is the oil, which is transparent in the adhesion map. Wetting properties of oil-rocks interfaces are important in industry because they are often used for oil reservoirs. Images c, d and e in Figure 3 show 3 force-distance curves, measured at 3 different points along a random diameter across the droplet used for analysis, as one can also see from image a. The more one goes towards the centre of the droplet, the higher adhesion and the higher trace and retrace forces are. The points of intersection between the height axis and trace and retrace forces are used as a droplet lower bound and droplet upper bound, respectively. Thus, the height of the droplet is calculated by subtracting the lower bound from the upper bound. The profile in Figure 4 describes the variation of the height along the chosen diameter. We notice the difference of contact angles between one end of the diameter and the other: 21° and 24°. This is associated with roughness along the contact line. Our study starts from this finding and focuses on developing a method to calculate interface properties which reflect the accuracy of AFM-imaged droplets in describing the degree of deviation of the droplet from a perfect spherical cap and the degree of roughness along the contact line which causes it. An image processing method for the AFM data is

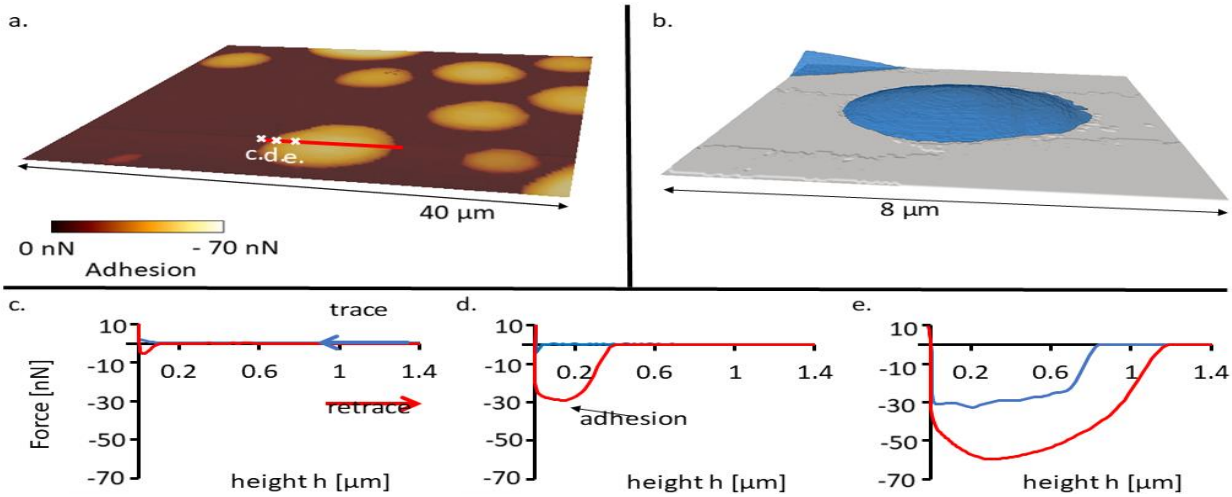


Figure 3 [5]:Adhesion map of a oil-water-calcite surface imaged with AFM (a.), the droplet of interest (b.) and 3 force distance curves along the diameter of the analysed droplet (c,d,e)

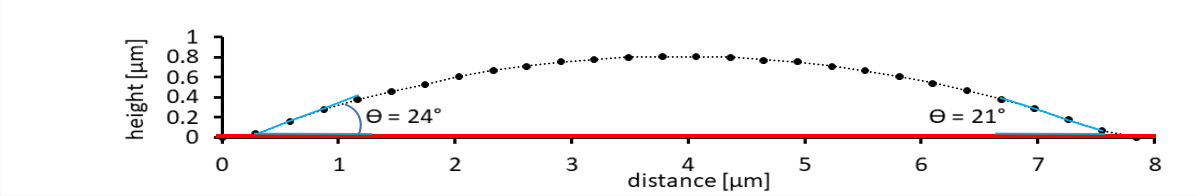


Figure 4 [5]:Height distance profile along the diameter of the droplet illustrated in Figure 1.a

developed and its sensitivity and efficiency for the defined purpose are characterized throughout this study.

2.Methodology

2.1 Voxelization

After the samples were imaged, JPKSPM Data Processing software was used to collect the numerical data, which characterizes the image. These data were then transformed into voxelized images, which were analysed by computational methods. Triangulation was the alternative to voxelization, but it uses approximations that reduce the level of accuracy in preserving the roughness of the original AFM data. Voxelization was therefore chosen and images with different resolutions were created in order to compare the accuracy of the results. The images analysed in this study have 3 resolutions: 512x512x512, 768x768x768 and 1024x1024x1024. The size of the original image is 37.175 µm on all three dimensions. Voxel sizes are therefore 37.175 /512 µm, 37.175/768 µm and 37.175/1024 µm

Data was collected from:

- the topographical image, which illustrates the position of the droplets in the scanned surface
- the adhesion image, which reveals the height profile along the solid surface
- the force distance curves, which show the height profile of the droplet, as it was explained in Section 1.

All these data were processed with an algorithm implemented in Matlab [Supplementary Material, 1] and transformed into a 3-phase 3D voxelized image, where voxels representing oil, water and calcite are labelled with 0, 1 and 2, respectively.

2.2 Principles of interface identification in a voxelized image

Volume of the droplet was firstly calculated by determining the number of voxels labelled with 1. The development of a script to identify 2-phase interfaces was the next step of this study. Figure 5 illustrates a sample 2D section through the 3D voxelized image, which highlights voxels representing 1 and 0, along with their interface. The orange line represents the real interface between 1 and 0, while the green line represents an approximation of the orange line, using voxelized area. In the first approach, a Matlab code was designed to identify the interface voxels and count them. The interface was visualised using Image J. The voxels corresponding to the interface were considered the ones whose value is 1 and whose diagonal or lateral neighbour is 0. In figure 5, these are the voxels surrounded by red and black squares. Because voxel resolution is 1x1x1 and it has a cubic shape, the values are the same in module for the volume of a cube voxel, the area of the face and the length of an edge. Thus, when more voxels are counted, the number of voxels represents in module all volume, area and length. Counting the interface voxels surrounded by squares is a measure for the area, corresponding in this case to the green line. Nonetheless, the green line is only an approximation of the real interface. The orange line describes the interface area with a higher accuracy, because the real interface does not

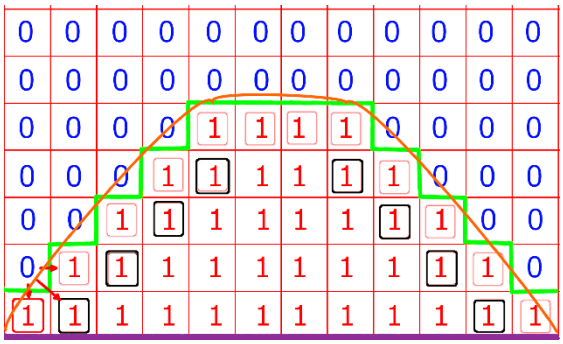


Figure 5: Sample 2D slice through the 3D matrix- the orange line represents the real interface between the two phases, while the green line represents the voxelized interface, which is equal to the number of voxels surrounded by squares

have a shape defined by cube voxels, as the green line does. The real interface has a curved shape, as the orange line does, because the outer interface of the droplet is expected to follow a quasispherical pattern. Therefore, a second improved method to calculate interface areas is developed, using Avizo 9.3.0.

2.3 Principles of interface identification in Avizo

Figure 6 presents the principles of an Avizo script [Supplementary Material, 4], whose target is to identify the solid-droplet interface (2-1 interface). Image 1 represents a sample 4x6 2D section through the 3D voxel image. The aim was to find the voxels surrounded by the green ellipse, which represent the solid-droplet interface in this sample slice. Two arithmetic functions were applied on the initial voxel slice shown in Image 1. The first one is a condition which checks if the value of the voxel is 1 and returns 0 or 1, if the condition is false or true. One can see in image 2 that the voxels representing exclusively the droplet remained 1, while all the other voxels turned to 0. The second function includes 2 conditions. The first one checks if the value of the voxel is 0, while the second one checks if the value of the voxel is 1. They were multiplied by 2 and added, so that the new image 3 separates the solid from both the liquids: 0 and 1 become 2, 2 becomes 0. "Label Interface" function was then applied on image 3 and the interface between liquid (both oil and droplet) and solid was determined. In order to obtain the droplet-solid interface, the all liquid-solid interface in image 4 was multiplied with only droplet voxels in image 2, so that voxels representing the oil-solid interface became 0, while voxels representing the droplet-solid interface remained 1. Therefore, image 5 illustrates only the droplet-solid interface voxels with values of 1. A similar algorithm, described by figure 7, was

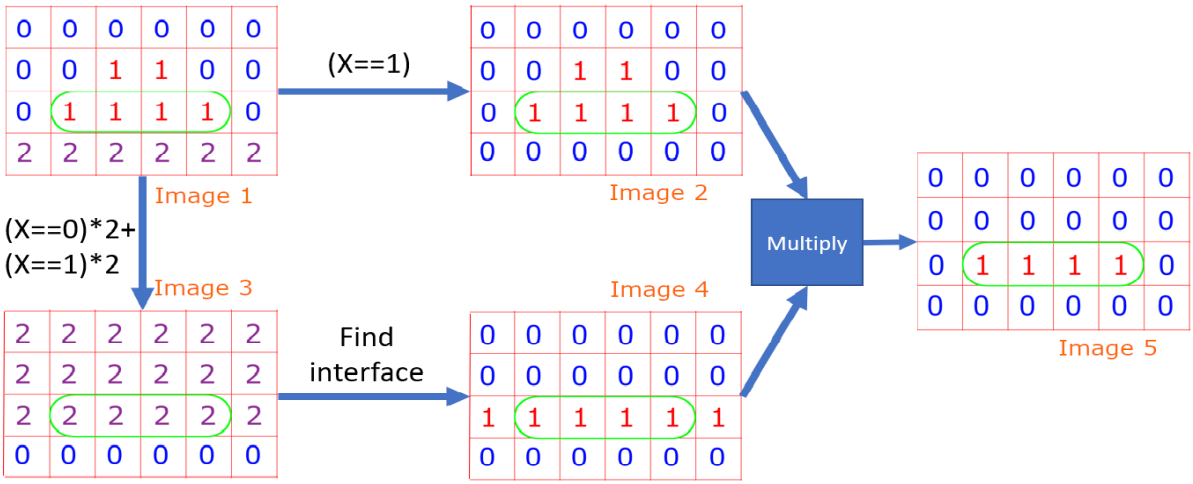


Figure 6: Algorithm to calculate interface between voxels representing 2(liquid) and 1(droplet) in Avizo

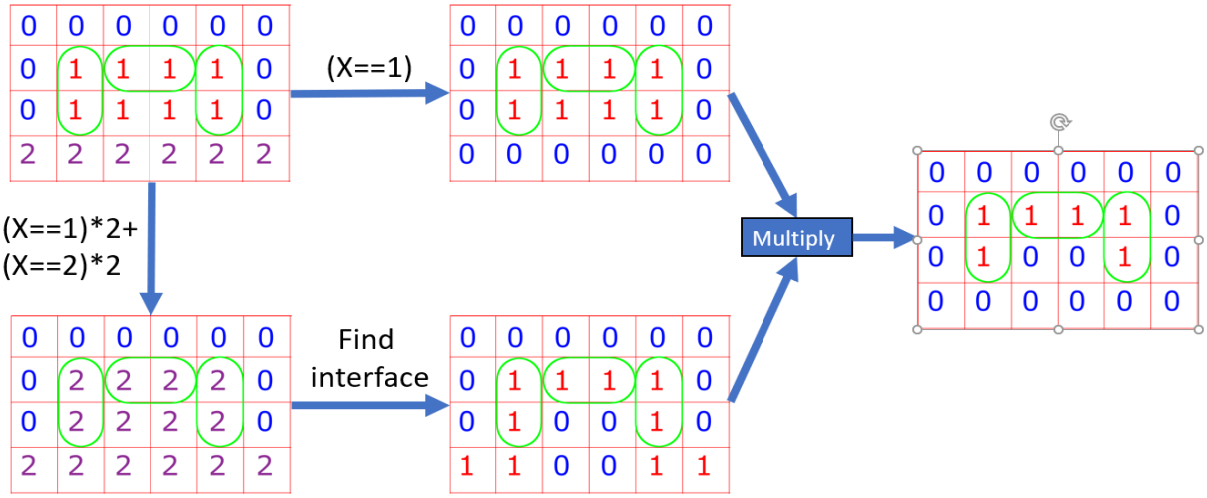


Figure 7: Algorithm to calculate the interface between the voxels representing 1(droplet) and 0(solid) in Avizo

used for creating the droplet-oil interface (interface between 2 and 1).

2.4 Development of perfect spherical cap

After identifying volume and 2-phase interface areas,a perfect spherical cap following a perfect sphere model with the same volume and base area of the real droplet was created in Matlab and deviation from the perfect spherical model was determined. The liquid-liquid interface for the perfect spherical cap was computed using the equations below [4], where a represents the radius of the base of the cap and h the height of the cap (as it is shown in Figure 8).

$$Droplet - Solid = Base\ Area = \pi a^2$$
$$Volume = \frac{1}{6}\pi h(3a^2 + h^2)$$
$$Liq. - Liq. = Curved\ Area = \pi(a^2 + h^2)$$

The contact angles of the perfect spherical caps were calculated using the equations below [4], where h is the height of the cap, r is the radius of the sphere and θ is the contact angle:

$$Liq - Liq\ \ Area = 2\,\pi rh = 2\pi r^2(1 - \cos\theta)$$

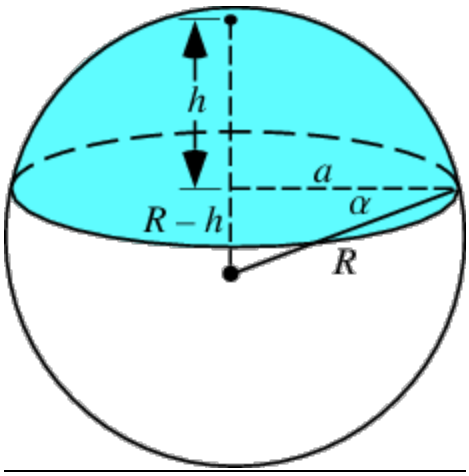


Figure 8: An example of the constructed spherical cap (the blue figure) [3]

2.5 The determination of the 3-phase contact line

The first approach for the identification of the 3-phase contact line and the consequent calculation of its length was similar to the approach for 2-phase interfaces as a principle. Therefore, it focused on developing a Matlab code that identifies if 3 consecutive voxels have the value of 2, 1 and 0 and count the voxels labelled with 1. As an alternative with higher accuracy, Avizo was used again. A build-in function which counts the average interface between the 3 phases was applied to the 3D voxelized image. For the 3-phase contact line, further cleaning and adjustment was undertaken. This is an essential step in cases where pinning, noise and high interface sensitivity occur.

2.6 The determination of the contact angle distribution

The voxelized image data was resampled into perpendicular 2D slices to the contact line and then measured automatically by tracing tangents to the solid surface and the oil–water interface, using the Scanziani-et-al contact angle measurement code ([3]). A contact angle distribution, including values measured at about 600-950 points on the contact line was thus determined.

The Matlab and Avizo codes and scripts that are used can be found in the Supplementary Material attached to this paper. Sensitivity of the software functions used for image processing was also investigated, following the results.

3. Results and Discussion

3.1 Volume

The volume of the selected droplet is 22.5 μm^3 for all resolutions with both methods of calculation. The error between the Avizo and Matlab code method is 0% in this case, because both algorithms calculate the number of voxels labelled with 1 for the same images. (Figure 9)

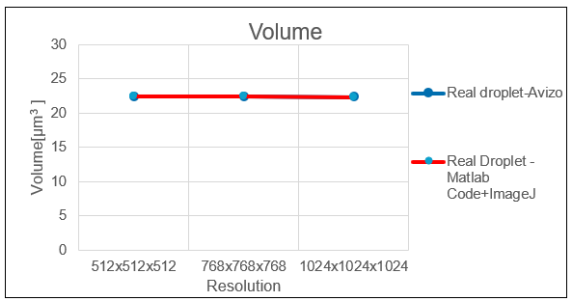


Figure 9: Volume as a function of the resolution- no dependency on resoluton

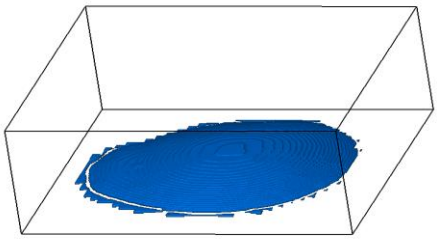


Figure 10: 3D visualization of the volume of the droplet- outer rim represents an image artefact

3.2 Droplet-solid interface

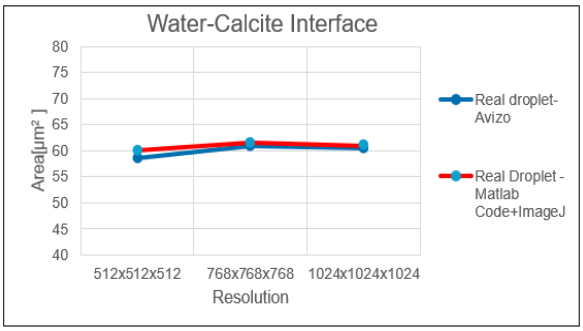


Figure 11: Water-calcite interface area as a function of resolution

The water-calcite interface has values in the 58.6-61.0 μm^2 range (Figure 11), which represents a maximum of 4.1% error. In contrast to the volume measurement, the water-calcite interface measurement has a degree of error, associated with voxelization and interface identification methods. The error between the Avizo and Matlab code methods is lower than 2.7% in all cases. The rest of the error can be attributed to the variation in resolution, which impacts the accuracy with which interfaces are measured. For example, one voxel from the 512x512x512 image becomes 2x2x2=8 voxels for the 1024x1024x1024 image. Therefore, there are cases where one voxel labelled with 0 in the 512x512x512 image becomes 4, 5, 6 or 7 voxels labelled with 0 and 4, 3, 2 or 1 voxels labelled with 1, respectively in the 1024x1024x1024 image, respectively. A voxel representing 0 with 100% in 512x512x512 image can thus become a group of voxels representing 0 with 50%, 62.5%, 75% or 87.5% in the 1024x1024x1024 image. The latter reflects a higher sensitivity of the interface. This fact illustrates that the accuracy of the information in creating interfaces is higher at higher resolutions and explains the variation obtained for the water-calcite interface at different resolutions.

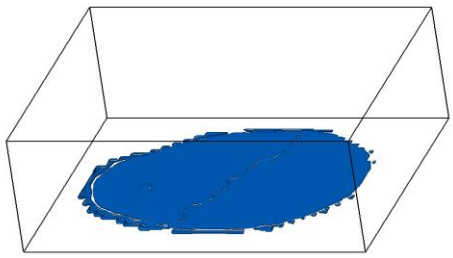


Figure 12: 3D visualization of the water-calcite interface

3.3 Comparison between the perfect spherical cap and the real droplet

The volume and the water calcite-interface are considered high precision measurements due to

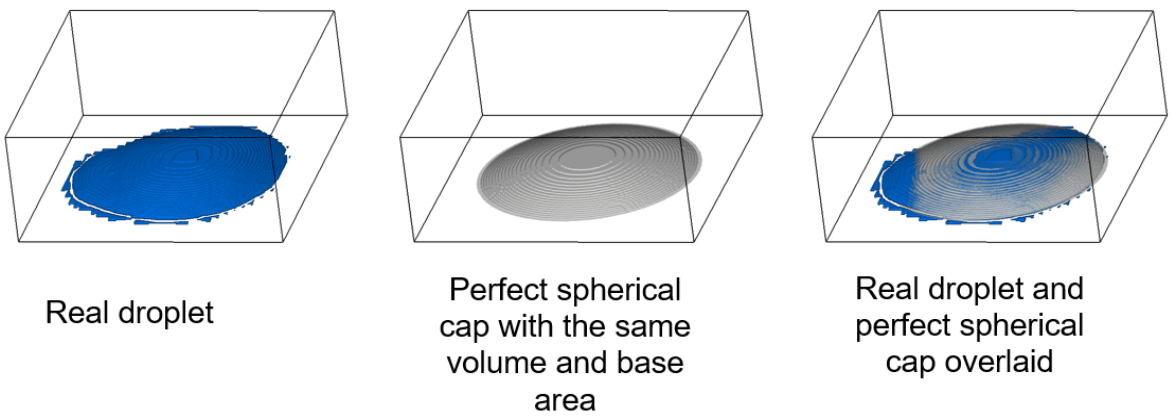


Figure 13: A comparison between the real droplet and the perfect spherical cap with the same volume and base area

low error between the 2 methods of measurement (0-2.7%). Therefore, a perfect spherical cap with the same volume and base area as the real droplet was developped in Matlab and compared to the real droplet. The two droplets were overlaid and aligned in figure 13. Although most of the voxels of the two droplets perfectly overlay, there is deviation from the perfect sphere model on the outer surface of the droplet, which represents the liquid-liquid interface. This can be attributed to the irregularity in nano-scale roughness along the contact line which creates a different alignment of the droplet along the contact line at each point and therefore generates the deviation from the perfect spherical cap model. If the solid had been perfectly flat and no roughness occurred, a perfect sphere model would have been expected, as calcite has a uniform surface from a chemical perspective. Along the contact line, pinning associated with the AFM measurement, voxelization process and surface roughness can also be observed. This is partly eliminated using the "Remove Small Spots" function in Avizo. However, although the pinning at the contact line has little effect on the interface calculation, it generates a low degree of error in all measurements which needs to be considered and further investigated in the future.

3.4 Liquid-liquid interface

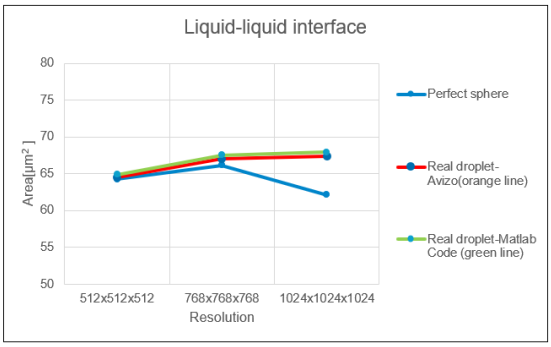


Figure 14: Liquid-liquid interface area as a function of resolution for the perfect sphere model and for the real droplet

In figure 14, the perfect spherical cap values are compared to the ones for real droplets analysed in both Avizo and with the Matlab code. As expected, the real droplet liquid-liquid interface obtained in Matlab is higher than the one in Avizo, which confirms the fact that Avizo identifies the curved averaged interface, rather than the voxelized interface (as shown in figure 5 and explained in Section 2.2). Also, both real values are higher than the perfect sphere model values. This is a geometrical consequence of the deviation from the perfect sphere. Moreover, the higher the resolution is, the higher the deviation is, with a maximum of 8.5% error for the highest resolution. This reinforces the fact that the voxelized image describes a more real pattern for higher resolutions.

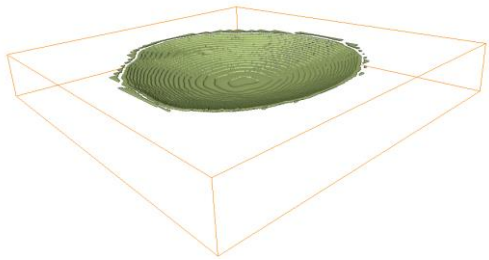


Figure 15: 3D visualization of the liquid-liquid interface with the droplet flipped upside down- this is the only view from which this interface can be seen

3.5 Contact Angle Distribution

3.5.1 Properties of the distribution

Contact angles were calculated for 602,940 and 911 points along the contact line for the 512x512x512, 768x768x768 and 1024x1024x1024 images, respectively. All 3 measurements correspond to a lower than 75 nm sampling interval along the contact line. Figure 16 illustrates the contact angle distribution for the 600-1000 points along the contact line. All contact angles range between 5° and 35°, while the average value for all the points on the contact line ranges between 19.04° and 20.66°. This represents a maximum error of 8.5% error for the 3 measured resolutions. The distribution takes a quasisymmetric pattern: the highest frequencies for the contact angles are around the average value, 20°. The frequencies decrease quasilinearly from 20° to both ends of the range, 5° and 35°. Each contact angle reveals a different level of roughness and therefore, this distribution illustrates that the level of the roughness on the calcite surface is different along the contact line, but it follows a pattern of

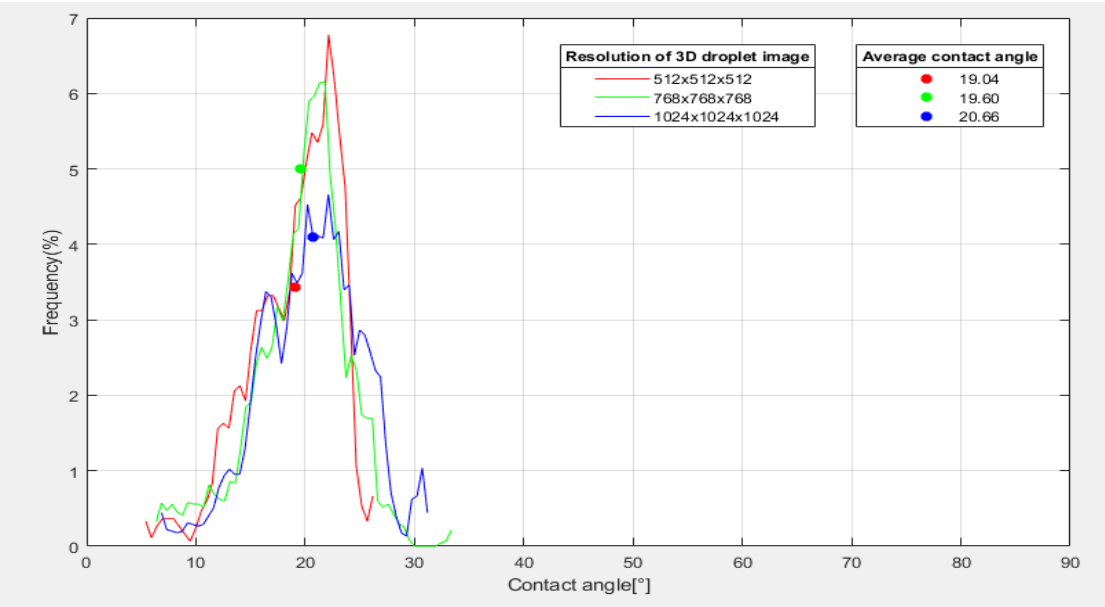


Figure 16: The distribution of the contact angles across the 3-phase contact line for the 3 resolution images

convergence towards its average level, corresponding to the average contact angle. The distribution of the contact angle demonstrates the roughness of the surface. The variation in contact angle could be also attributed to factors such as inclination in the AFM probe or chemical heterogeneity, but these are excluded in this study. If the AFM probe was not perfectly flat and that was the cause for the deviation of the contact angles, the distribution could not be quasisymmetric with a convex shape; it would be linear. Also, the calcite surface has low chemical heterogeneity, but if other solids are used, this aspect needs to be investigated for each case.

3.5.2 Effects of resolution

The contact angle distribution demonstrates that accuracy and resolution are proportional. The higher the resolution is, the wider the distribution range of the contact angles is. The lower resolution images with less voxels enclose the same information as the higher resolution images with more voxels. Therefore, the contact angles values obtained for lower resolutions are measured on more approximated surfaces than the contact angles for higher resolutions. This explains the increase in the distribution range of the contact angles for higher resolutions: for 1024x1024x1024, the range is 6.5°-32.0°, while for 512x512x512, it is 5.0° -26.5°.

Furthermore, this is also reinforced by the change in the highest peaks in frequency. At higher resolution, where values are spread over a wider range, the highest peak is lower than for lower resolutions, where contact angles are spread over a narrower range, with more averaged values. One can observe this variation in Figure 16, where the highest peaks for 512x512x512, 768x768x768 and 1024x1024x1024 resolutions are at 6.8%, 6.1% and 4.7%, respectively.

3.5.3 Deviation from the perfect spherical cap

The average contact angles of the real droplet were compared to the contact angles of the perfect spherical cap. Figure 17 shows the deviation between the two models ranges between 0 and 8.4% and increases with the resolution. The distortion of the real droplet from the perfect model can thus be characterized.

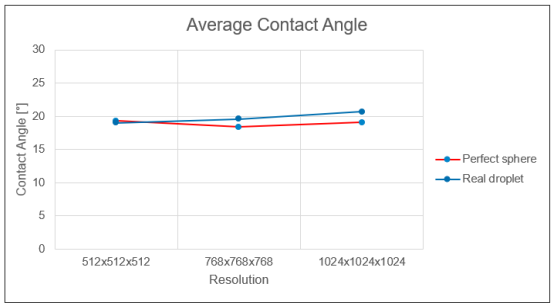


Figure 17: Comparison between the perfect sphere contact angle and the average contact angle along the contact line

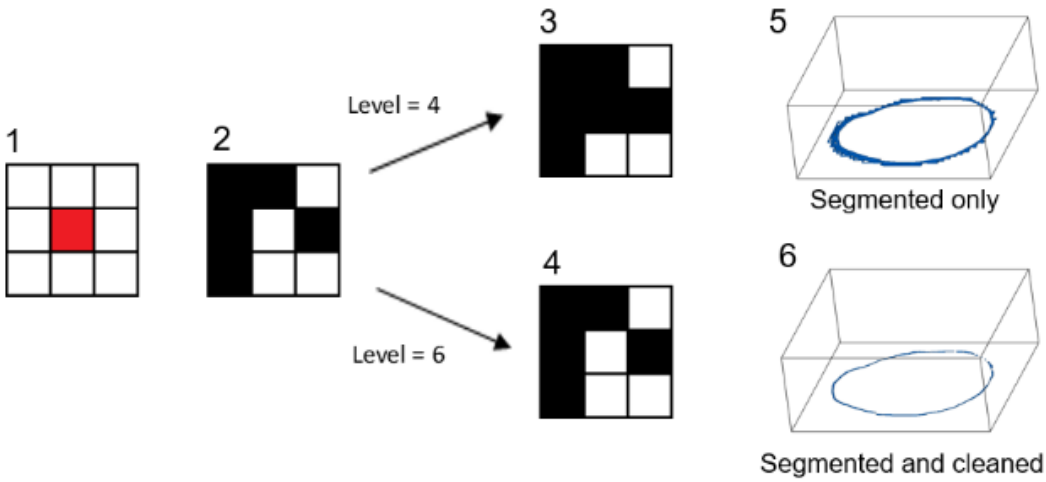


Figure 18:Analysis of voxel filtering sensitivity(1,2,3,4) [5] and 3D visualizations of segmented only 3-phase contact line(5) and of segmented and cleaned 3-phase contact line(6)

3.6 Three-phase contact line

In what regards the 3-phase contact line, voxel filtering sensitivity is crucial. In Figure 18, the principles of the filtering in Avizo are pointed out. The red voxel in image 1 is surrounded by 5 black voxels in image 2. The aim is to identify if this is an interface voxel between the black and the white interfaces. If the level of sensitivity is higher than the number of surrounding black voxels (which is 5 in this case), the voxel will be identified as an interface voxel, as image 3 illustrates. If not, it will remain a white voxel and image 4 will be obtained. Avizo only has 3 levels of sensitivity regarding voxel filtering: 6, 18 and 26. This corresponds to values determined geometrically: 6, 18 and 26 are the numbers of cube voxels with one face in common with the analysed voxel, at least one edge in common with the analysed voxel and at least one corner in common with the analysed voxel, respectively. Avizo cannot crop the voxels or make a statistical analysis and therefore, the voxel filtering has this low sensitivity, which impacts the identification of the 3-phase contact line. The most accurate contact line available in Avizo with the interface labelling function for the measured water droplet can be seen in Figure 18.5. This was further cleaned and transformed into a more realistic contact line, presented in Figure 18.6. However, the precision of the cleaning process is low and the level of error still remains high. The 3-phase contact line cannot have 2-4 voxels in thickness, as our identified contact line has.

This suggests that significant error in measurement appear. This statement is further consolidated by the graph in Figure 19. The contact line in the segmented only image has a length, which is 4 times higher than the perfect sphere, while the line in the segmented and cleaned image still has a length which is about 2.2 times higher than the perfect sphere. The error is partly caused by the deviation of the real contact line from a perfect circle, but this cannot justify a 120-300% error. Evidently, most of the deviation is caused by the low sensitivity of the voxel filtering function in Avizo.

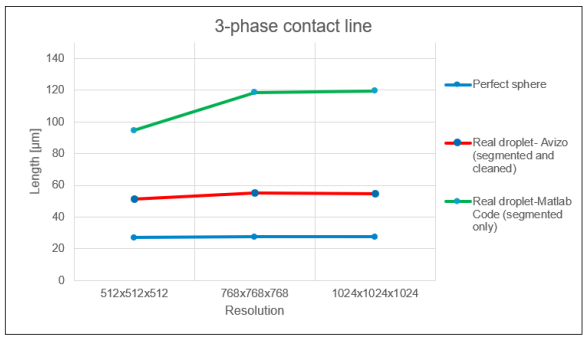


Figure 19: 3-phase contact line length as a function of resolution

4.Conclusion

This study demonstrates that the accuracy of contact angle measurements of AFM-imaged droplets is higher than for micro-CT, because this method takes roughness into consideration accurately and uses images scanned at a smaller scale. However, pinning and noise around the

contact line still occur because of experimental measurement and voxelization. Contact angle, 2-phase interface areas and volumes can be calculated with a high accuracy using this method and their values illustrate the range of roughness across the droplet-solid interface and 3-phase contact line. The length of the latter cannot be calculated accurately, because of the contrast between low voxel filtering sensitivity and high contact line sensitivity for the voxelized image. Also, the analysis on resolution proved that the best accuracy of the voxelization method is achieved at the highest resolution. Consequently, when measurements are repeated using this algorithm, highest available resolution is recommended. This study can be used as a start point for characterizing wetting through the level of roughness surface and across 3-phase contact lines.

5.Outlook

In the future, the method to calculate interfaces and contact angles should be investigated for more droplets for both oil-water-calcite surfaces and other composition surfaces. This technique can be applied on fluid film measurements from rocks [5]. Sensitivity of this method as a function of the composition and shape of the surface is an important next step. Also, a new software or code needs to be developed for the calculation of the length of the 3-phase contact line. Methods to eliminate pinning and noise when scanning with atomic force microscopy and during the voxelization process require further investigation.

6.References

[1] David Quéré (2008), *Wetting and Roughness,* Paris: Annual Review of Materials Research, Vol. 38, 71-99. Available online at: https://www.annualreviews.org/doi/10.1146/annurev.matsci.38.060407.132434

[2] Matthew Andrew, Branko Bijeljic, Martin J. Blunt (2014), *Pore-scale contact angle measurements at reservoir conditions using X-ray microtomography*, London: Elsevier, Advances in Water Resources, Vol. 68, 24-31. Available online at: https://doi.org/10.1016/j.advwatres.2014.02.014

[3] Eric Weinstein. *Spherical Cap*, MathWorld-A Wolfram Web Resource. Available online at: http://mathworld.wolfram.com/SphericalCap.html

[4] Alessio Scanziani, Kamajit Singh, Martin J. Blunt, Alberto Guadagnini (2017), *Automatic method for estimation of in situ effective contact angle from X-ray micro tomography images of two-phase flow in porous media,* Journal of Colloid and Interface Science, Vol. 496, 51-59. Available online at: https://doi.org/10.1016/j.jcis.2017.02.005

[5] Maja Rücker, Willem-Bart Bartels, Tom Bultreys, Marijn Boone, Kamaljit Singh, Gaetano Garfi, Alessio Scanziani, Catherine Spurin, Sherifat Yesufu-Rufai, Samuel. Krevor, Martin. J. Blunt, Ove Wilson, Hassan Mahani (2019), *Workflow for upscaling wettability from the nano- to core-scales*, Pau, France: International Symposium of the Society of Core Analysts. Available online at:https://www.researchgate.net/publication/335442553_Workflow_for_upscaling_wettability_from_the_nano-to_core-scales

Towards Practical Application of Real-Time Optimisation using Gaussian Processes

Guan Huang and Titus Ong
Department of Chemical Engineering, Imperial College London, South Kensington, London SW7 2AZ, U.K.

Abstract. Real-time optimisation (RTO) leverages on online plant measurements to deal with plant-model mismatch and disturbances. This paper builds upon the recent work by del Rio Chanona et al. (2019) by investigating the practical performance of modifier-adaptation (MA) schemes utilising machine learning through Gaussian processes (GP). These GPs are built using steady-state data and avoid plant gradient estimation in classic MA schemes. Several practical challenges such as having a more complex system, increased input space, and stricter restrictions on constraint violations have been embedded into an Imperial College Carbon Capture Pilot Plant optimisation problem. We also present two excitation sub-algorithms which improve the predictions made with GP. The efficacy of this RTO approach is illustrated with case studies and this is one step closer to actual implementation of the algorithm in real-life.

1 Introduction

Process optimisation is key to maximising the performance of a plant while ensuring its operational safety. Traditional offline optimisation has improved plant operations through modelling; however, they neither address process disturbances well nor reach the optimum if the model is inaccurate (Marchetti et al., 2016). Real-Time Optimisation (RTO) methods have been developed over the past 30 years to address these issues by incorporating live process measurements into the optimisation process (Bonvin, 2017). Even so, popular methods such as the "two-step" approach, where process measurements are used to re-estimate model parameters iteratively (Chen and Joseph, 1987), often do not converge at plant optimum due to differing KKT conditions arising from structural plant-model mismatch (Tatjewski, 2002; Chachuat et al., 2009).

One RTO method that has been proven to deal with plant-model mismatch is the Modifier Adaptation (MA) method (Marchetti et al., 2009). Instead of updating model parameters, the MA method updates first-order modifiers in the cost and constraint functions of the model which correct the error between plant and model. This drives the model toward plant optimum as the model KKT conditions match those of the plant upon convergence (Chachuat et al., 2009). However, MA is heavily dependent on plant gradient estimation which can be costly due to additional plant evaluations and difficult to estimate with noisy data (Marchetti et al., 2009).

Recently, Ferreira et al. (2018) proposed the novel use of Gaussian Processes (GP) in MA schemes. GP is a probabilistic regression technique that has been extensively developed over the past few decades in the field of machine learning (Williams and Rasmussen, 2006). It is popular because of its ease for inference and learning over other machine learning methods and for its intuitive model predictions. In Ferreira et al.'s use case, the classic MA modifiers are replaced with predictors of GPs that are trained recursively with steady-state process measurements. As the GPs captured the unknown system behaviour well, Ferreira et al.'s algorithm outperformed MA in noise attenuation and had the added benefit of avoiding plant gradient estimation.

del Rio Chanona et al. (2019) further developed the MA-GP algorithm by introducing trust region ideas, also known as restricted steps method, in a manner similar to input filtering in MA. This restricts the step size taken by the system, hence ensures that extrapolation of the GPs are under control and plant constraint violations are reduced.

Nonetheless, these algorithms were only tested on a Williams-Otto reactor benchmark problem, hence have not been subjected to tests on a scale of a process plant. Some challenges that will arise include a larger input space, a more complex optimisation problem, and a more careful approach required to handle constraint violations which may pose safety hazards.

This paper aims to develop the MA-GP and trust region algorithm for practical application, with the end goal of applying it on the Imperial College Carbon Capture Pilot Plant (IC CCPP). Modifications are made to this algorithm to address the aforementioned practical challenges, and two additional sub-algorithms for excitation are considered. The rest of the paper provides background on MA, GP and trust region in Section 2, describes the modified algorithm in Section 3, illustrates its performance in Section 4, and presents the excitation sub-algorithms in Section 5.

2 Preliminaries

2.1 Problem Formulation

The steady-state optimisation of a plant can be formulated as the following optimisation problem:

$$\begin{aligned} \min_{\boldsymbol{u} \in \mathcal{U}} \quad & G_0^p(\boldsymbol{u}) := g_0(\boldsymbol{u}, \boldsymbol{y}^p(\boldsymbol{u})) \\ \text{s.t.} \quad & G_i^p(\boldsymbol{u}) := g_i(\boldsymbol{u}, \boldsymbol{y}^p(\boldsymbol{u})) \leq 0, i = 1 \dots n_g \end{aligned} \tag{1}$$

where $\boldsymbol{u} \in \mathbb{R}^{n_u}$ is the vector of input variables; $\boldsymbol{y}^p \in \mathbb{R}^{n_y}$ is the vector of output variables; $g_0 : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}$

is the cost function to be minimised; $g_i : \mathbb{R}^{n_u} \times \mathbb{R}^{n_y} \to \mathbb{R}, \quad i = 1...n_g$ are the inequality constraint functions; $\mathcal{U} \subseteq \mathbb{R}^{n_u}$ is the control domain, which is typically the bounds on the input variables, $\boldsymbol{u}^L \leq \boldsymbol{u} \leq \boldsymbol{u}^U$. The superscript $(\cdot)^p$ denotes a plant-related quantity.

However, the plant mapping $\boldsymbol{y}^p(\boldsymbol{u})$ is usually not known accurately and optimisation is done on the best approximate model $\boldsymbol{y}(\boldsymbol{u}, \boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$ is the vector of adjustable model parameters. Thus, the following model-based optimisation problem is solved instead:

$$\begin{aligned} \min_{\boldsymbol{u} \in \mathcal{U}} \quad & G_0(\boldsymbol{u}) := g_0(\boldsymbol{u}, \boldsymbol{y}(\boldsymbol{u}, \boldsymbol{\theta})) \\ \text{s.t.} \quad & G_i(\boldsymbol{u}) := g_i(\boldsymbol{u}, \boldsymbol{y}(\boldsymbol{u}, \boldsymbol{\theta})) \leq 0, i = 1 \dots n_g \end{aligned} \tag{2}$$

Due to plant-model mismatch and plant disturbances, the optimum of problem 2 often does not match that of problem 1. RTO deals with this by incorporating live process measurements to update the model in problem 2 and overcome process uncertainties, hence driving the model to converge at plant optimum.

2.2 Modifier Adaptation

In MA schemes, process measurements are used to update cost and constraint functions via first-order modifiers (Marchetti et al., 2009) while model parameters $\boldsymbol{\theta}$ are not updated. The MA-based problem at iteration k to solve for $\boldsymbol{u}^{k+1}$ can then be formulated as such:

$$\begin{aligned} \boldsymbol{u}^{k+1} \in \arg\min_{\boldsymbol{u} \in \mathcal{U}} \quad & G_0(\boldsymbol{u}) + (\boldsymbol{\lambda}_{G_0}^k)^T \boldsymbol{u} \\ \text{s.t.} \quad & \\ & G_i(\boldsymbol{u}) + \varepsilon_{G_i}^k + (\boldsymbol{\lambda}_{G_i}^k)^T [\boldsymbol{u} - \boldsymbol{u}^k] \leq 0, i = 1 \dots n_g \end{aligned} \tag{3}$$

where $\varepsilon_{G_i}^k \in \mathbb{R}$ are zeroth-order modifiers for the constraints, and $\boldsymbol{\lambda}_{G_i}^k \in \mathbb{R}^{n_u}$ are the first-order modifiers for the cost and constraints. These modifiers are updated using live measurements at iteration k. This ensures that the model reaches the same necessary optimality conditions as the plant upon convergence, i.e. a KKT point in problem 3 is a KKT point in problem 1.

2.3 Gaussian Processes

A GP describes an unknown function $f : \mathbb{R}^{n_u} \to \mathbb{R}$ given a set of inputs $\boldsymbol{U} := [\boldsymbol{u}^1 \dots \boldsymbol{u}^N] \in \mathbb{R}^{n_u \times N}$ and corresponding outputs $\boldsymbol{y} := [y^1 \dots y^N]^T \in \mathbb{R}^N$. Each observation may be noisy and can be represented by $y = f(\boldsymbol{u}) + \nu$, where $\nu \sim \mathcal{N}(0, \sigma_\nu^2)$ is Gaussian noise with variance σ_ν^2.

Then, a GP generalises multivariate Gaussian distributions, such that a subset of the range follows a multivariate Gaussian distribution for some domain (Ebden, 2015):

$$f(\boldsymbol{u}) \sim \mathcal{GP}(m(\boldsymbol{u}), k(\boldsymbol{u}, \boldsymbol{u'})) \tag{4}$$

where $m(\boldsymbol{u})$ is the mean function and $k(\boldsymbol{u}, \boldsymbol{u'})$ is the covariance function. For our purposes, we use a constant mean function (Eqn. 5) and a squared exponential covariance function (Eqn. 6), which was chosen because it approximates smooth functions well (Williams and Rasmussen, 2006):

$$m(\boldsymbol{u}) := c \tag{5}$$

$$k(\boldsymbol{u}, \boldsymbol{u'}) := \sigma_n^2 \exp\Big[\frac{-(\boldsymbol{u} - \boldsymbol{u'})^T \boldsymbol{\Omega} (\boldsymbol{u} - \boldsymbol{u'})}{2}\Big] \tag{6}$$

The hyperparameters present in Eqn. 6 are the scaling matrix $\boldsymbol{\Omega} := \text{diag}(\lambda_1 \dots \lambda_{n_u})$ and max covariance σ_n^2.

The unknown hyperparameters $\boldsymbol{\Psi} := [c \ \sigma_n \ \sigma_\nu \ \lambda_1 \dots \lambda_{n_u}]^T$ are commonly inferred using maximum likelihood estimation. The log-likelihood can be written as follows:

$$\mathcal{L}(\boldsymbol{\Psi}) := \frac{-\ln(|\boldsymbol{K}(\boldsymbol{U})|) - (\boldsymbol{y} - \boldsymbol{I}c)^T \boldsymbol{K}(\boldsymbol{U})^{-1} (\boldsymbol{y} - \boldsymbol{I}c)}{2} \tag{7}$$

with $K_{ij}(\boldsymbol{U}) := k(\boldsymbol{u}^i, \boldsymbol{u}^j) + \sigma_\nu^2 \delta_{ij}, \quad (i, j) \in \{1 \dots N\}^2$ where δ_{ij} is the Kronecker delta function.

We can then predict the distribution at a point $\boldsymbol{u}$:

$$f(\boldsymbol{u}) | (\boldsymbol{U}, \boldsymbol{y}) \sim \mathcal{N}(\mu_f(\boldsymbol{u}), \sigma_f^2(\boldsymbol{u})) \tag{8}$$

where predicted mean $\mu_f(\boldsymbol{u})$ and predicted variance $\sigma_f^2(\boldsymbol{u})$ are given by the following equations:

$$\mu_f(\boldsymbol{u}) := \boldsymbol{r}(\boldsymbol{u}, \boldsymbol{U}) \boldsymbol{K}(\boldsymbol{U})^{-1} \boldsymbol{y} + c \tag{9}$$

$$\sigma_f^2(\boldsymbol{u}) := \sigma_n^2 - \boldsymbol{r}(\boldsymbol{u}, \boldsymbol{U}) \boldsymbol{K}(\boldsymbol{U})^{-1} \boldsymbol{r}(\boldsymbol{u}, \boldsymbol{U})^T \tag{10}$$

$$\boldsymbol{r}(\boldsymbol{u}, \boldsymbol{U}) := [k(\boldsymbol{u}, \boldsymbol{u}^1) \dots k(\boldsymbol{u}, \boldsymbol{u}^N)] \tag{11}$$

2.4 Input trust region

Trust region methods are employed in various optimisation techniques and represents a subset of input space around the current operating point where the outputs can be predicted with reasonable accuracy (Conn et al., 2000). It can then shrink or expand in size depending on the level of confidence one has in the model.

Consider the MA-based problem described in problem 3. We can add a trust region constraint to limit the size of the step $\boldsymbol{d}^{k+1}$ to be taken from $\boldsymbol{u}^k$ to $\boldsymbol{u}^{k+1}$:

$$\boldsymbol{d}^{k+1} \leq \Delta^k \tag{12}$$

where Δ^k is the radius of the trust region.

We then define ρ^{k+1}, a ratio of the actual plant improvement to predicted improvement:

$$\rho^{k+1} = \frac{G_0^p(\boldsymbol{u}^k) - G_0^p(\boldsymbol{u}^{k+1})}{[G_0 + \mu_{\delta G_0}^k](\boldsymbol{u}^k) - [G_0 + \mu_{\delta G_0}^k](\boldsymbol{u}^{k+1})} \tag{13}$$

where $\rho^{k+1} \simeq 1$ means that the model prediction is fairly accurate, and $\rho^{k+1} \ll 1$ indicates that the plant performed worse than expected.

For decision making, ρ is often compared to user-defined parameters η_1 and η_2, where $0 \leq \eta_1 \leq \eta_2 \leq 1$. These thresholds determine whether the trust region should shrink or expand by a factor of t_1 or t_2 respectively. Lastly, there is choice for the initial Δ^0 and maximum Δ^{max} trust region radius as well.

3 Methodology

3.1 MA utilising GP modifiers

The use of GPs to capture plant-model mismatch and serve as higher-order modifiers was first proposed by Ferreira et al. (2018). GPs were trained and utilised for correcting cost and each constraint functions individually.

$$G_i^p - G_i \sim \mathcal{GP}(\mu_{\delta G_i}, \sigma^2_{\delta G_i}), \quad i = 0 \dots n_g \tag{14}$$

Eqn. 14 denotes the GP approximation trained with a set of data $(\boldsymbol{U}, \boldsymbol{\delta G}_i)$; $\boldsymbol{U} := [\boldsymbol{u}^1 \dots \boldsymbol{u}^N]$; $\delta G_i(\cdot) := G_i^p(\cdot) - G_i(\cdot)$ is the steady-state mismatch between the plant measurements and model outputs at each input point. The mean function $\mu_{\delta G_i}$, which is a GP predictor, is then used in the proposed MA with GP modifiers scheme (MA-GP) with the following modified optimisation problem:

$$\begin{aligned} \boldsymbol{u}^{k+1} \in \arg\min_{\boldsymbol{u} \in \mathcal{U}} \quad & [G_0 + \mu^k_{\delta G_0}](\boldsymbol{u}) \\ \text{s.t.} \quad & [G_i + \mu^k_{\delta G_i}](\boldsymbol{u}) \leq 0, \quad i = 1 \dots n_g \end{aligned} \tag{15}$$

3.2 MA-GP with input trust region algorithm

del Rio Chanona et al. (2019) extended the MA-GP optimisation problem 15 and its corresponding algorithm by introducing trust region ideas in two fashions. This paper will focus on the first variant where trust region is implemented in the input space. The modified optimisation problem is as follows:

$$\begin{aligned} \boldsymbol{d}^{k+1} \in \arg\min_{\boldsymbol{d}} \quad & [G_0 + \mu^k_{\delta G_0}](\boldsymbol{u}^k + \boldsymbol{d}) \\ \text{s.t.} \quad & [G_i + \mu^k_{\delta G_i}](\boldsymbol{u}^k + \boldsymbol{d}) \leq 0, \quad i = 1 \dots n_g \\ & \|\boldsymbol{d}\| \leq \Delta^k, \quad \boldsymbol{u}^k + \boldsymbol{d} \in \mathcal{U} \end{aligned} \tag{16}$$

where $\boldsymbol{d}^{k+1} \in \mathbb{R}^{n_u}$ is the predicted step.

The formalisation of the MA-GP with input trust region algorithm (MA-GP-ITR) can be found in Algorithm 2 of del Rio Chanona et al. (2019). The algorithm is initialised with a set of data that is used to build the first iteration of GPs. Suitable trust region parameters are also chosen for the problem at hand. During each iteration, the modified optimisation problem 16 is solved to find the optimal point within the current trust region. The next courses of action—(1) to move or backtrack the operating point (2) how to adjust the trust region radius—is then decided by decision matrices. If any constraints are violated, the trust region radius Δ^{k+1} shrinks and the operating point $\boldsymbol{u}^{k+1}$ remains the same in the next iteration. This is essentially backtracking in practical RTO. Otherwise, the trust region radius and the operating point are updated depending on the value of ρ^{k+1}. In the last step, the data set used for training is updated with information from the new step and the GPs improve as the iteration progresses.

Ferreira et al. (2018) and del Rio Chanona et al. (2019) have both presented case studies based on the Williams-Otto reactor with an initial data set of 5 training points which span across a considerable input space and concluded the following:

1. A lower number of plant evaluations were needed to reach plant optimum as plant gradient estimations were no longer needed
2. GP-based MA outperforms standard MA in terms of handling noisy data
3. The implementation of trust regions ensures extrapolation of GPs remains under control and thus improves reliability
4. Trust regions increased convergence speed

3.3 Modifications to MA-GP-ITR for practical RTO

Algorithm 1 Modified practical MA-GP-ITR

Initialisation: Input (1) initial data sets $(\boldsymbol{U}^0, \boldsymbol{\delta G}_i^0)$ where matrix $\boldsymbol{U}^0$ consists of points $\boldsymbol{u}$ within the initial trust region (2) initial feasible operating point $\boldsymbol{u}^0 \in \mathcal{U}$ (3) trust region parameters $0 < \eta_1 < \eta_2 \leq 1$, $0 < \Delta^0 < \Delta^{max}$, and $0 < t_1 < 1 < t_2$.
Repeat: for $k = 0, 1, \dots$

(1) Train GP predictors $\mu^k_{G_i}$, $i = 1 \dots n_g$ with normalised data sets $(\boldsymbol{U'}^k, \boldsymbol{\delta G}_i^k)$

(2) Solve the modified optimisation problem 16 to obtain $\boldsymbol{u}^{k+1} := \boldsymbol{u}^k + \boldsymbol{d}^{k+1}$

(3) Obtain measurements of the cost and constraint functions $G_i^p(\boldsymbol{u}^{k+1})$, $i = 1 \dots n_g$

(4) Compute ρ^{k+1} according to Eqn. 13

(5) Update the trust region radius:

- **If** $G_i^p(\boldsymbol{u}^{k+1}) > 0$ for any $i = 1 \dots n_g$ **or** $\rho^{k+1} < \eta_1$:
  $\Delta^{k+1} := t_1 \Delta^k$
- **Else if** $\rho^{k+1} > \eta_2$ and $\left\|\boldsymbol{d}^{k+1}\right\| = \Delta^k$:
  $\Delta^{k+1} := min\{t_2 \Delta^k, \Delta^{max}\}$
- **Else** :
  $\Delta^{k+1} := \Delta^k$

(6) Update the operating point

- **If** $G_i^p(\boldsymbol{u}^{k+1}) > 0$ for any $i = 1 \dots n_g$ **or** $\rho^{k+1} < \eta_1$:
  $\boldsymbol{u}^{k+1} := \boldsymbol{u}^k$
- **Else**:
  $\boldsymbol{u}^{k+1} := \boldsymbol{u}^k + \boldsymbol{d}^{k+1}$

(7) Update the data sets $(\boldsymbol{U}^{k+1}, \boldsymbol{\delta G}_i^{k+1})$ with measurements at $\boldsymbol{u}^k + \boldsymbol{d}^{k+1}$ **if** $G_i^p(\boldsymbol{u}^{k+1}) < 0$ for all $i = 1 \dots n_g$

In this section, slight modifications made to the MA-GP-ITR algorithm (by del Rio Chanona et al.) to evaluate its performance and behaviour in a practical application are presented. The full statement of the modified algorithm is provided in **Algorithm 1** and the MATLAB code can be found in App. B. Firstly, the input variables of the data sets used to train GPs are always normalised against its operating range:

$$\boldsymbol{u}' = \frac{\boldsymbol{u} - 0.5(\boldsymbol{u}^U + \boldsymbol{u}^L)}{0.5(\boldsymbol{u}^U - \boldsymbol{u}^L)} \tag{17}$$

Normalisation is performed so that each input variable will contribute proportionately to its normalised range as it is usual to have input variables on vastly different scales. This allows the covariance matrix built with Eqn. 6, which calculates Euclidean distance, to have minimal bias towards an input variable which has a different scale.

Notice that the initial training points are chosen from only within the initial trust region. This is to reflect situations in real life when historical data may be difficult to retrieve or when the plant in question has only ever been operating in a narrow input space. This poses a challenge as the first GPs may not be as accurate and is more susceptible to data noise.

In the final step of the original MA-GP-ITR algorithm, the data sets are always updated with new information at the end of each iteration. In a simulation environment, these measurements are gathered at $\boldsymbol{u}^{k+1}$ even when constraints are violated as there are no implications in a computational context. However, in practical RTO, if the physical plant is moving towards a new set point $\boldsymbol{u}^{k+1}$ and approaches a constraint, controller or operator intervention will prevent the plant from actually reaching $\boldsymbol{u}^{k+1}$. This is especially in cases when these constraints are safety constraints or stringent product target constraints. Thus, no new information is gathered and we modify the final step by including a "no constraint violation" condition for the data set update. This increases the difficulty of the RTO as exact positions of the constraints are difficult to determine and the algorithm relies heavily on the trust region parameters.

4 Case Studies

With the end goal of applying the RTO scheme on the IC CCPP, an Aspen HYSYS V9 flowsheet of the physical system has been developed for use in the case studies in this paper (see App. A for more details). The IC CCPP utilises 15% monoethanolamine (MEA) solution which is looped between an absorber column, where it captures CO_2 from an inlet flue gas stream of only CO_2 and N_2, and a stripper column where the solvent is regenerated through heating. To model this, the rate-based Aspen HYSYS acid gas package is used and pump electrical usage is fitted to historical data. The reader is directed to Wang et al. (2017) for more details on carbon capture technology; Dyment and Watanasiri (2019) for reaction set used in Aspen HYSYS acid gas package and model validation; Hale (2017), Ganapathy and Jacob (2013, 2014) and Clapson and Dong (2017) for details of the IC CCPP which are used in the simulation.

Algorithm 1 was implemented in MATLAB and a HYSYS-MATLAB link was set up to control the HYSYS simulation. At each iteration, *fitrgp* of the Statistics and Machine Learning Toolbox was used to generate GPs and NLP solver *fmincon* of the Optimization Toolbox was used to solve the optimisation problem at hand in similar fashion to del Rio Chanona et al. (2019).

Table 1. Trust region parameters for input variables

	Q_{st}/kJ h^{-1}	F_{recyc}/kg h^{-1}	F_{inlet}/kg h^{-1}
Δ^0	5,500	50	15
Δ^{max}	1,100	100	30

To build up case studies for RTO application, 2 input variables were chosen: (1) the reboiler duty of the stripper column Q_{st} (2) the solvent recycle flowrate F_{recyc}. These are the 2 main inputs used in both case studies 1 (Section 4.1) and 2 (Section 4.2) that directly affect the costs of operation and carbon capture performance. An addition input variable, flue gas inlet flowrate F_{inlet}, is used in case study 2. Their respective trust region parameters are chosen conservatively where a full step with maximum trust region radius will not break the physical plant. The values are given in Table 1. Similarly, conservative values of $t_1 = 0.8$, $t_2 = 1.2$ and common trust region parameter settings of $\eta_1 = 0.1$, $\eta_2 = 0.9$ are used. All these parameters are used throughout this section and in Section 5.

This paper presents a model-free approach where $G_i = 0, \quad i = 0 \ldots n_g$. Thus, Eqn. 14 is updated as $G_i^p \sim \mathcal{GP}(\mu_{G_i}, \sigma^2_{G_i}), \quad i = 0 \ldots n_g$ where the GP is now capturing the full plant behaviour and not just the plant-model mismatch. Eqns. 13 and 15 are updated accordingly as well. No doubt this poses an even greater challenge but reflects situations where plant models may not be available at hand or the problem is so complex that models are of low accuracy. That said, a low-fidelity model could have been used in this work as an approximate model to the HYSYS model to introduce "plant-model mismatch".

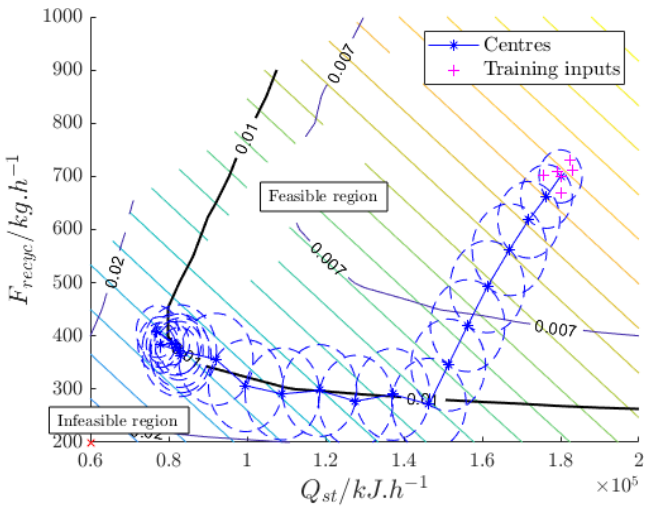


Figure 1. RTO of problem 18. The contours of the objective function is shown in thin multicoloured lines while the purity inequality constraint is represented by the thick black line. The RTO iterates are plotted with a blue line with dashed blue circles representing trust regions.

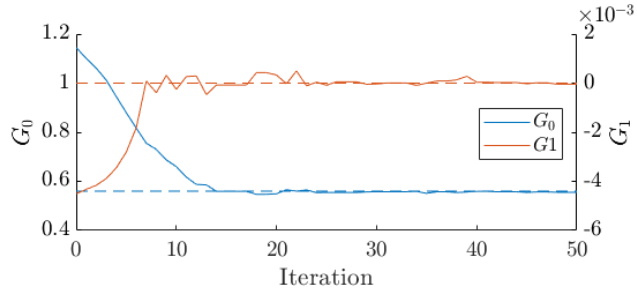
4.1 Case study 1: Operational optimisation

Consider the optimisation problem which aims to minimise the operation cost given an inlet flue gas flowrate F_{inlet} of $80\,\mathrm{kg\,h^{-1}}$ with inlet CO_2 composition X_{inlet,CO_2} of 10 vol% while subject to an outlet gas stream purity constraint X_{outlet,CO_2} of 1 vol%:

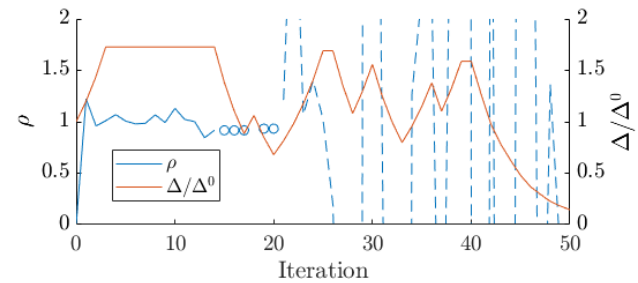
$$\begin{aligned}
\min_{Q_{st},F_{recyc}} \quad & G_0 := 0.009F_{st} + 0.25Q_{elec} \\
\text{s.t.} \quad & F_{inlet} = 80 \\
& X_{inlet,CO_2} = 0.10 \\
& F_{st} = f(Q_{st}) \\
& Q_{elec} = f(Q_{st}, F_{recyc}) \\
& X_{outlet,CO_2} = f(Q_{st}, F_{recyc}) \\
& G_1 := x_{outlet,CO_2} - 0.01 \leq 0 \\
& Q_{st} \in [60,000,\ 200,000] \\
& F_{recyc} \in [200,\ 1000]
\end{aligned} \tag{18}$$

The objective function G_0 is constructed as the total price of utilities where F_{st} is the steam flowrate and Q_{elec} is the electricity usage of the process pumps. F_{st}, Q_{elec} and X_{outlet,CO_2} are given by the HYSYS model and thermodynamic relations. The cost G_0 and inequality constraint G_1 measurements are assumed to have zero mean noise with standard deviations $\sigma_{G_0} = 0.0025$ and $\sigma_{G_1} = 0.00004$.

Fig. 1 depicts the RTO graphically in the input space where the unconstrained optimum is in the bottom left corner of the plot and Fig. 2 shows the performance of the RTO algorithm. Finally, Fig. 3 shows the evolution of the GPs with increasing iterations.

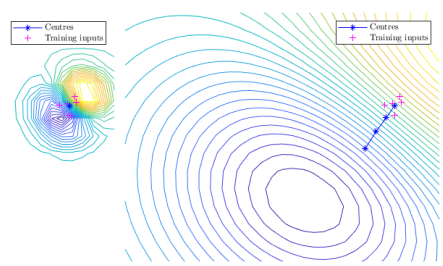


(a) G_0, G_1 against iteration number

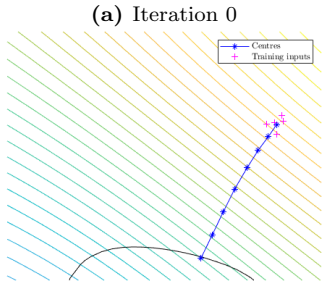


(b) Parameters ρ, Δ against iteration number

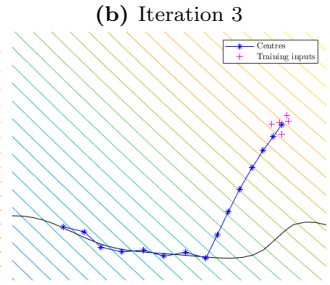
Figure 2. Operational optimisation plots. **(a)** Dashed blue line: plant optimum value. Dashed orange line: value of zero. **(b)** Blue circles: Iterations where constraint is violated and no ρ value is computed. Dashed blue line: Erratic ρ values which are out of range of interest.



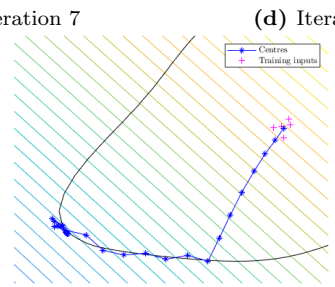
(a) Iteration 0 **(b)** Iteration 3



(c) Iteration 7



(d) Iteration 14



(e) Iteration 50

Figure 3. Development of GPs for cost function G_0 (contours of multicoloured thin lines) and inequality constraint G_1 (black line)

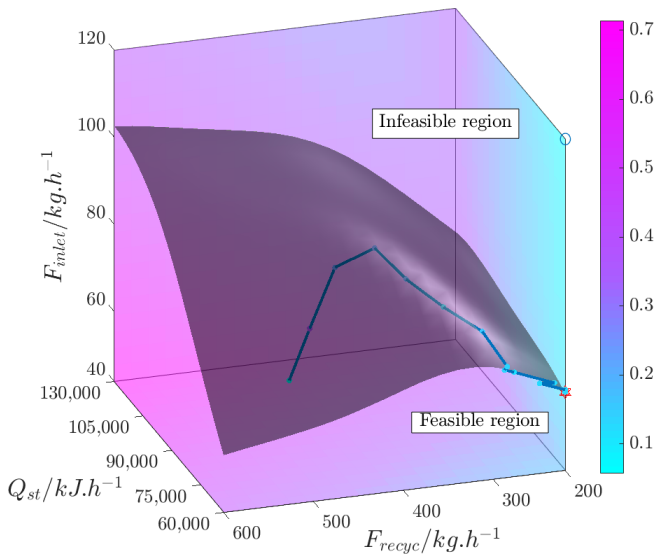


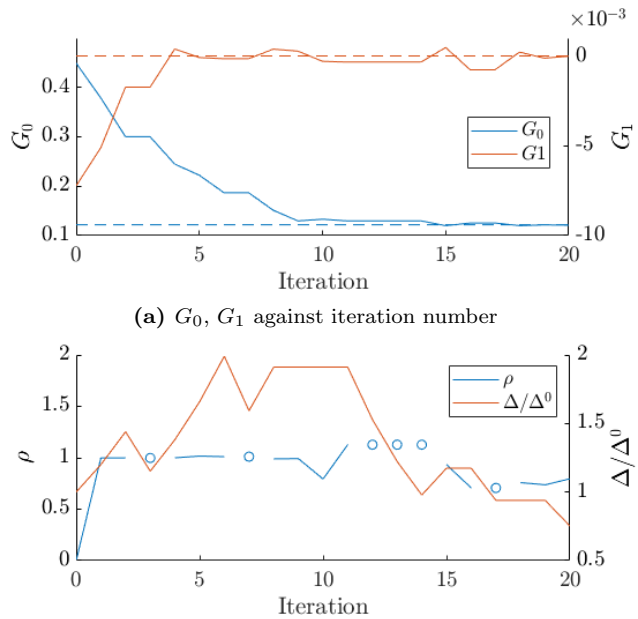
Figure 4. RTO of problem 19. The value of the objective function is represented in colour while the purity inequality constraint is represented by the 3D black surface. The RTO iterates are plotted with a blue line with the starting point being the left-most point.

The algorithm moves the plant inputs towards the 1% purity constraint quickly before travelling along it and converging at the constrained plant optimum. This happens within 15 iterations despite a rather restrictive trust region size. The value of ρ is close to 1 for these iterations and shows the accuracy of the GP predictors despite noise and being purely data-driven.

We can understand the optimiser's behaviour by analysing the GP predictions (Fig. 3). With the initial training inputs (Fig. 3a), predictions are messy due to the close proximity of the points. The range of objective values is narrow and highly susceptible to noise. However, the algorithm can still move the plant in the right direction and, as more data is gathered with an expanding objective value range, the GP predictions become more accurate and resembles the actual plant behaviour.

In iterations 15-20, the constraint is violated several times due to the GP prediction mismatch (see ρ values in Fig. 2b). Fig. 3d shows that the GP has not captured the sharp turn feature of the constraint and thus the optimiser naturally wants to continue on a leftward trajectory. The trust region size shrinks accordingly in response to the violation and is able to keep further iterates within the feasible region. Nonetheless, this provides a strong motivation for a combined model-data approach since having an approximate idea of the sharp turn would reduce the number of constraint violations.

Near the plant optimum, the steps taken are small and changes in the objective value becomes insignificant. This causes the value of ρ to become erratic and meaningless. After about 40 iterations, the algorithm converges at one spot and the delta shrinks continuously to close the loop.



(a) G_0, G_1 against iteration number

(b) Parameters ρ, Δ against iteration number

Figure 5. Economic optimisation plots. **(a)** Dashed blue line: plant optimum value. Dashed orange line: value of zero. **(b)** Blue circles: Iterations where constraint is violated and no ρ value is computed.

Note that a small numerical tolerance is included for constraint violation. Though it seems that some iterates have ventured into the infeasible region in Fig. 1, Fig. 2 shows that the value of the violation is insignificant.

4.2 Case study 2: Economic optimisation

This section presents another case study from an economic optimisation point of view. The input space is extended to 3 variables where $\boldsymbol{u}$ consists of $Q_{st}, F_{recyc}, F_{inlet}$ and the optimisation problem is as follows:

$$
\begin{aligned}
\min_{\boldsymbol{u}} \quad & G_0 := 0.009F_{st} + 0.25Q_{elec} \\
& \quad - 31F_{CO_2\,captured} \\
\text{s.t.} \quad & x_{inlet,CO_2} = 0.10 \\
& F_{st} = f(Q_{st}), \quad Q_{elec} = f(Q_{st}, F_{recyc}) \\
& F_{CO_2\,captured} = f(Q_{st}, F_{recyc}, F_{inlet}) \\
& x_{outlet,CO_2} = f(Q_{st}, F_{recyc}, F_{inlet}) \\
& G_1 := x_{outlet,CO_2} - 0.01 \leq 0 \\
& Q_{st} \in [60,000,\ 130,000] \\
& F_{recyc} \in [200,\ 600] \\
& F_{inlet} \in [40,\ 120]
\end{aligned}
\tag{19}
$$

The objection function now includes an economic benefit term of a hypothetical carbon tax avoided through the operation of the plant. The added decision variable F_{inlet} assumes that there is control over the inlet flue gas flowrate and this may be achieved by venting excess or by having a pre-treatment storage unit. In essence, in order to reap greater economic benefits by treating a

larger volume of gas, utilities usage has to increase as well to meet the outlet purity constraint.

A graphical representation of the RTO is shown in Fig. 4 where the unconstrained plant optimum (blue circle) and the constrained optimum (red hexagram) are plotted. The first few steps are in the direction towards the unconstrained plant optimum until the constraint is met. The optimiser then changes direction swiftly to move down along the constraint until it hits the input lower bound for reboiler duty before converging at the plant optimum.

The conservative trust region parameters chosen contributed to the rather sudden change in direction after meeting the constraint. A larger trust region size may allow the RTO to take the hypotenuse path but the trade off would be loss of accuracy and possibly more steps that violate the constraint.

Despite the increase in number of decision variables and a more complex constraint surface, **Algorithm 1** is still able to perform well and predict plant behaviour and constraints accurately as shown in Fig. 5.

In contrast to the classic MA where the problem becomes substantially more challenging with the increase of input variables, the use of GP seems to be able to circumvent this difficulty as there is no need to carry out $(n_u + 1)$ perturbations at each iteration for gradient estimation. The method presented in this paper should be able to deal with even bigger optimisation problems fairly well and there still exists considerable flexibility in choosing the form of the covariance function for GPs, trust region parameters and the inclusion of a base model.

5 Excitation

5.1 Motivation

The GP models give fairly accurate estimates, as demonstrated in the previous section, even with a small and local training data set as mentioned in Section 3.3. However, this can lead to "blind spots" in the GP where the training data set does not adequately cover the output surfaces. As a result, minimising the GP of the cost function can produce a step in a sub-optimal direction.

We have identified two symptoms where this occurs: firstly, when the GP takes parallel steps because it has little information about the orthogonal plane; secondly, when the GP takes small or negligible steps and converges at a sub-optimal point because it has little information about the local conditions.

The premise of excitation is taking a step that is different from the step given by solving the optimisation problem. Hence, we explore excitation during these two conditions to improve the GP model, not as a way to find gradients but to feed data into the GP that it would not have obtained otherwise to increase its accuracy from a machine learning point of view. The excitation sub-algorithms are described in Section 5.2 and exemplified in Section 5.3. The MATLAB code is available at App. B.

5.2 Methodology

5.2.1 Generalised excitation sub-algorithm

Sub-algorithm 1 Excitation

Step 1 of **Algorithm 1** is performed.

(1) Check if excitation condition is met:
- **If** true:

  Go to step 2 of **Sub-algorithm 1**
- **Else**:

  Return to step 2 of **Algorithm 1**

(2) Find excited point $\boldsymbol{u}^{k+1} := \boldsymbol{u}^k + \boldsymbol{d}^{k+1}$

(3) Obtain measurements of the cost and constraint functions $G_i^p(\boldsymbol{u}^{k+1}), \quad i = 1 \dots n_g$

(4) Update the trust region radius:
- **If** $G_i^p(\boldsymbol{u}^{k+1}) > 0$ for any $i = 1 \dots n_g$:

  $\Delta^{k+1} := t_1 \Delta^k$
- **Else** :

  $\Delta^{k+1} := \Delta^k$

(5) Update the operating point
- **If** $G_i^p(\boldsymbol{u}^{k+1}) > 0$ for any $i = 1 \dots n_g$:

  $\boldsymbol{u}^{k+1} := \boldsymbol{u}^k$
- **Else If** $G_0^p(\boldsymbol{u}^{k+1}) > G_0^p(\boldsymbol{u}^k)$:

  $\boldsymbol{u}^{k+1} := \boldsymbol{u}^k$
- **Else**:

  $\boldsymbol{u}^{k+1} := \boldsymbol{u}^k + \boldsymbol{d}^{k+1}$

Return to step 7 of **Algorithm 1**.

The excitation sub-algorithms can be generalised into Sub-algorithm 1, and this can be fitted in Algorithm 1 detailed in Section 3.3. The condition is checked in step 1 of the sub-algorithm and if the condition is met, the excitation algorithm is run before returning to step 7 of Algorithm 1.

The key differences between Algorithm 1 and Sub-algorithm 1 are the decision matrices for updating the trust region radius and operating point in steps 4 and 5. Excited points do not have valid ρ values as they can be a worse step compared to the current operating point, while ρ calculation pre-supposes an improving step. Hence, the trust region radius that depends on ρ is kept unchanged except for constraint violation as it is only relevant for taking a predicted step. Likewise for

the operating point, the excited point cannot be evaluated against ρ and is thus evaluated against its cost function value.

The following two sections describe steps 1 and 2 and expand on the two sub-algorithms separately.

5.2.2 Parallel steps excitation

The system has taken two steps in parallel if it fulfils the following condition:

$$1 - \frac{(\boldsymbol{u}^k - \boldsymbol{u}^{k-1})}{\left\|(\boldsymbol{u}^k - \boldsymbol{u}^{k-1})\right\|} \cdot \frac{(\boldsymbol{u}^{k-1} - \boldsymbol{u}^{k-2})}{\left\|\boldsymbol{u}^{k-1} - \boldsymbol{u}^{k-2}\right\|} < \varepsilon_p \tag{20}$$

where ε_p is a numerical tolerance.

We define a "GP feasibility" constraint where a point does not violate the GP-modelled constraints:

$$\mu^k_{G_i}(\boldsymbol{u}^{k+1}) \leq 0, \quad i = 1 \ldots n_g \tag{21}$$

Then, the excited point $\boldsymbol{u}^{k+1}$ is the furthest point in the orthogonal direction that is subject to GP feasibility and trust region constraints (Eqn. 12). We take a far step to minimise the effect of noise and give the GPs a better overall picture of the cost and constraint surfaces. This can be formulated as the following equations:

$$\boldsymbol{u}^{k+1} := \boldsymbol{u}^k + x\hat{\boldsymbol{\nu}}^k \tag{22}$$

$$\hat{\boldsymbol{\nu}}^k \cdot (\boldsymbol{u}^k - \boldsymbol{u}^{k-1}) = 0 \tag{23}$$

where x is the maximum value such that the aforementioned constraints are not violated.

5.2.3 Negligible steps excitation

The past two steps are considered negligible within a tolerance ε_n if they fulfil the following condition:

$$\left\|\boldsymbol{u}^{k-i} - \boldsymbol{u}^{k-i-1}\right\| \leq \varepsilon_n, \quad i = 0, 1 \tag{24}$$

Then, the excited point $\boldsymbol{u}^{k+1}$ is a randomly generated point subject to GP feasibility and trust region constraints (Eqns. 12,21).

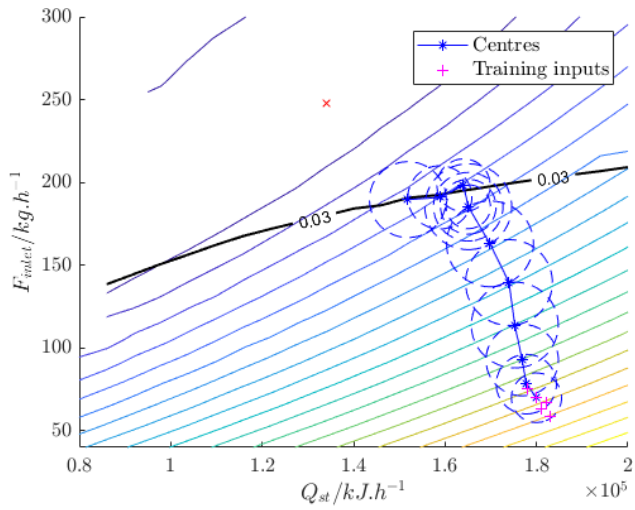
5.3 Case Studies

Parameter values ε_p = 1e-7, ε_n = 5e-2 were used.

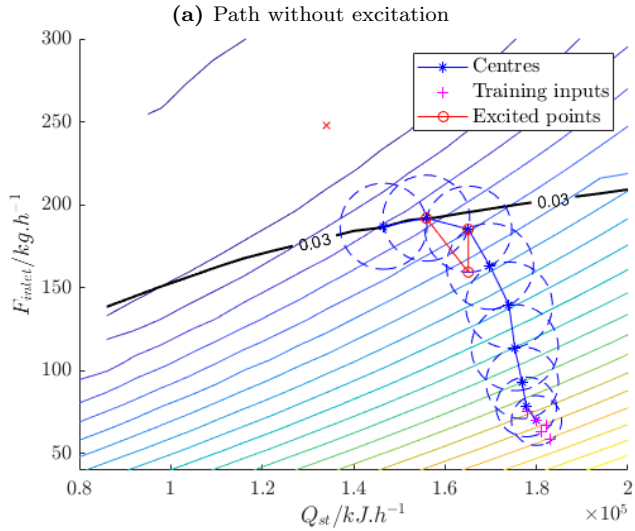
5.3.1 Case study 3: Parallel steps excitation

The optimisation problem for this case study is based on the economic optimisation problem in Section 4.2 but with a 3% purity constraint and without electrical costs in the cost function. The input space has been reduced to 2 dimensions for clearer visualisation of the results, so $F_{recyc} = 3 \times F_{inlet}$ for a constant G/L ratio.

Fig. 6a depicts the path of optimisation upon meeting the constraint without parallel steps excitation, while Fig. 6b presents the path taken with excitation. Without excitation, the G_1 constraint is violated thrice as evidenced by the shrinking of the trust region thrice. In contrast, the optimiser navigates along the constraint curve without any violations by just having one excitation.



(a) Path without excitation

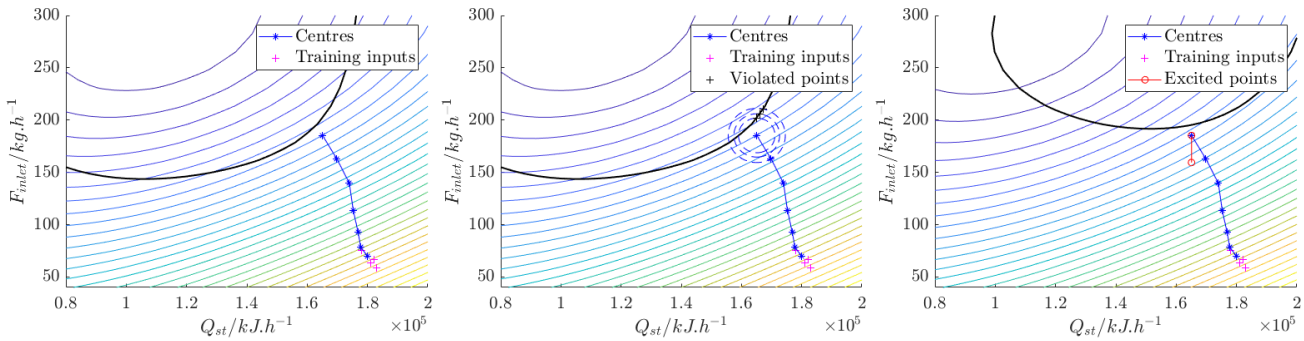


(b) Path with excitation

Figure 6. Case study 3 contour plots (thin multicoloured lines) of HYSYS model cost function. Thick black line: G_1 constraint. Dashed blue circles: Trust regions. Blue line: Path of optimisation. Red cross: HYSYS optimum.

An examination of the GP contour plots in Fig. 7 shows the effectiveness of excitation. Initially at iteration 6 in Fig. 7a, the optimiser has been moving mostly in a straight line and has poor information on the constraint curve at its sides. As such, it thinks that the optimal step is in the right direction of the constraint curve and keeps violating G_1 until the trust region is small enough to restrict the step effectively in Fig. 7b. With excitation, the GP has a better idea of the constraint curve and the optimiser moves left to avoid constraint violation.

It is evident that an additional data point in the orthogonal plane increases the accuracy of the GP's predictions for the cost and constraint functions. With only a limited training data set, the GP is prone to continually predict a step in a similar direction especially as it moves further away from its starting point. Because



(a) Iteration 6 **(b)** Iteration 7 without excitation **(c)** Iteration 7 with excitation

Figure 7. Case study 3 contour plots (thin multicoloured lines) of GP-modelled cost function at various iterations. Thick black line: GP G_1 constraint. Dashed blue circles: Trust regions. Blue line: Path of optimisation.

the new data points are in the same plane, each iteration of GP gets more information about this plane and can cause a positive feedback loop of stepping in the same direction. Excitation in the orthogonal plane gives the GP a better sense of the input space, enabling it to step in the correct direction and reduce constraint violations. Nevertheless, this is at the cost of an additional step and the optimiser may be biased towards the side of excitation.

Overall, parallel steps excitation makes exploration safer and faster where the cost and constraint planes are difficult to traverse, so the benefits suggest that excitation is worth implementing.

5.3.2 Case study 4: Negligible steps excitation

The optimisation problem for this case study is similar to that in Case Study 3 (Section 5.3.1) but inclusive of the electrical costs. Fig. 8 depicts the contour plot for this problem. Note that the optimum is not bounded by constraints but is along the Q_{st} input bounds.

Fig. 9a shows that the optimiser does not converge at the plant optimum without any excitation. Even with parallel steps excitation, the GP optimum is still along the constraint (Fig. 9b). With negligible steps excitation in Fig. 9c, the optimiser moves downwards until it reaches the plant optimum.

A comparison between Fig. 9a and Fig. 9b once again presents the case for parallel steps excitation, since the latter GP has a better estimation of the cost plane. Nonetheless, we note that this excitation is best at providing information across the input planes and not for finding local information. In comparison, the randomness of negligible steps excitation allows the optimiser to explore the local region and confirm its optimality. However, this also leads to inconsistency and the optimiser can take a long time to find the local optimum.

Since the negligible steps excitation sub-algorithm would kick in when the optimiser starts converging at a point, majority of the optimisation path is not affected so this excitation is feasible in confirming local optimality at little cost.

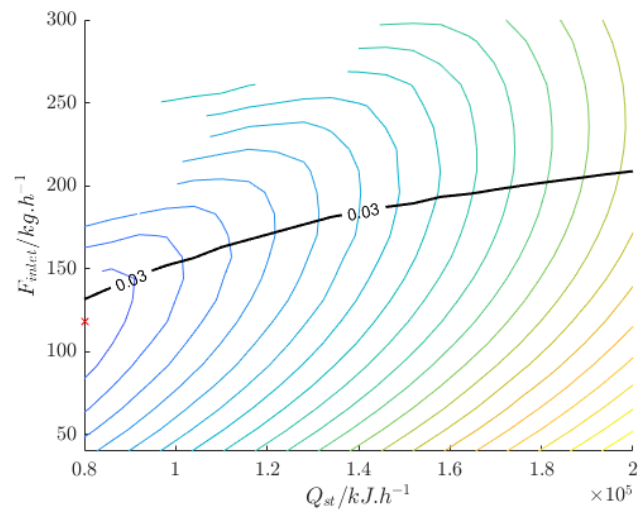
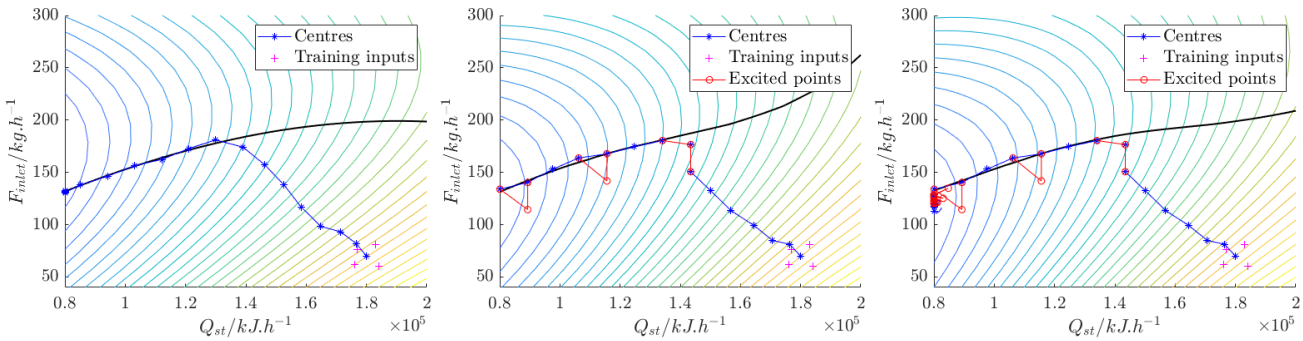


Figure 8. Case study 4 contour plot (thin multicoloured lines) of HYSYS model cost function. Thick black line: G_1 constraint. Red cross: HYSYS optimum.

6 Conclusion and Future Work

This paper builds on previous work done by del Rio Chanona et al. (2019) by modifying the MA-GP-ITR scheme for practical application. The modified algorithm was successfully demonstrated to account for practical challenges such as large input space and harsher constraint violation restrictions. Despite using a model-free approach, the RTO converged at the optimum in both the operational and economic optimisation problems. It was noted that a combined model and data-driven approach might present better results and this will be investigated in the future plant implementation. Two excitation sub-algorithms were also explored under the conditions of parallel steps and negligible steps and were shown to be feasible in making GP predictions more robust with few disadvantages.

To further develop the algorithm to respond to constraint violations, the use of transient measurements will be considered. Dynamic model identification may be able to extract information from these measurements and allow the algorithm to forecast and avoid further



(a) No excitation **(b)** Only parallel steps excitation **(c)** Parallel and negligible steps excitation

Figure 9. Case study 4 contour plots (thin multicoloured lines) of GP-modelled cost function with various sub-algorithms. Thick black line: GP G_1 constraint. Blue line: Path of optimisation.

constraint violations. A framework for excitation has been developed but there is much room to improve its synergistic use with trust regions.

With these results, we have more confidence that this algorithm can be implemented in the IC CCPP, although several challenges will have to be addressed before actual implementation. The connection between the algorithm in MATLAB and the IC CCPP Control System has to be bridged and methods for steady-state identification has to be tested. The present HYSYS model will be used as a base model for increased reliability and faster convergence speed.

7 Bibliography

Bonvin, D. (2017), 'Special Issue "Real-Time Optimization" of Processes', *Processes* **5**(4), 27.

Chachuat, B., Srinivasan, B. and Bonvin, D. (2009), 'Adaptation strategies for real-time optimization', *Computers Chemical Engineering* **33**(10), 1557 – 1567. Selected Papers from the 18th European Symposium on Computer Aided Process Engineering (ESCAPE-18).

Chen, C. and Joseph, B. (1987), 'On-line optimization using a two-phase approach: An application study', *Ind. Eng. Chem. Res.; (United States)* **26**.

Clapson, L. and Dong, J. (2017), 'Viability of simulator for use in operator training and model based control on a carbon capture process'.

Conn, A. R., Gould, N. I. M. and Toint, P. L. (2000), *Trust Region Methods*.

del Rio Chanona, E., Graciano, J. A., Bradford, E. and Chachuat, B. (2019), 'Modifier-adaptation schemes employing gaussian processes and trust regions for real-time optimization', *IFAC-PapersOnLine* **52**(1), 52 – 57.

Dyment, J. and Watanasiri, S. (2019), 'Acid gas cleaning using amine solvents: Validation with experimental and plant data'.

Ebden, M. (2015), 'Gaussian processes: A quick introduction'.

Ferreira, T. D. A., Shukla, H. A., Faulwasser, T., Jones, C. N. and Bonvin, D. (2018), Real-Time optimization of Uncertain Process Systems via Modifier Adaptation and Gaussian Processes, *in* '2018 European Control Conference, ECC 2018'.

Ganapathy, J. and Jacob, Y. (2013), 'Expanding the operating envelope of an amine solvent carbon capture pilot plant - interim report'.

Ganapathy, J. and Jacob, Y. (2014), 'Expanding the operating envelope of an amine solvent carbon capture pilot plant'.

Hale, C. P. (2017), 'Pilot Plant: Carbon Capture Process'.

Marchetti, A., Chachuat, B. and Bonvin, D. (2009), 'Modifier-adaptation methodology for real-time optimization', *Industrial & Engineering Chemistry Research* **48**(13), 6022–6033.

Marchetti, A. G., François, G., Faulwasser, T. and Bonvin, D. (2016), 'Modifier adaptation for real-time optimization—methods and applications', *Processes* **4**(4), 55.

Tatjewski, P. (2002), 'Iterative optimizing set-point control – the basic principle redesigned', *IFAC Proceedings Volumes* **35**(1), 49 – 54. 15th IFAC World Congress.

Wang, Y., Zhao, L., Otto, A., Robinius, M. and Stolten, D. (2017), 'A review of post-combustion co2 capture technologies from coal-fired power plants', *Energy Procedia* **114**, 650 – 665. 13th International Conference on Greenhouse Gas Control Technologies, GHGT-13, 14-18 November 2016, Lausanne, Switzerland.

Williams, C. K. and Rasmussen, C. E. (2006), *Gaussian processes for machine learning*, Vol. 2, MIT press Cambridge, MA.

Designing polyurethane foams aimed at sportswear market

Iylia Ismail and Swee Kee Hoo
Department of Chemical Engineering, Imperial College London, U.K.

Abstract This study takes an innovative approach in developing polyurethane foam (PUF) for fashionable sportswear as a protective element. External protective elements such as shin pads and helmets are considered old-fashioned in the sportswear industry because of their bulkiness. A silicone-based polymer (SP) integrated into fabric was recently developed by a fashion designer, Kiren Passi. This incorporation of material into fabric reduces the necessity of external protective elements. However, sagging issue occurred in the fabric due to the material being relatively heavy. Thus, PUF was suggested as a replacement to the silicone-based polymer. Preliminary test results showed that densities of PUF was 83% less than SP in average, which assured avoidance of the sagging complication. Formulations developed had various PUF properties, and hence gave distinct result in moulding and mechanical integrity. Examples of formulations are: rigid foam (BDO,TDI)(1:1.05) and soft foam (PEG 600,TDI)(1:1.9). Rigid foams' compressive modulus was approximately 40 times higher than SP, and they took the shape of a deep mould whereas soft PUFs performed slightly worse in these aspects. However, soft PUFs have low deformation indicator values as they withstood temperature ranges from 37.5°C to 60°C. All formulations have common issues with irregular sizes and dispersion of bubbles, giving an appearance which was less attractive.

Introduction

Polyurethane was first discovered by Professor Dr. Otto Bayer in 1930s and it has been a major breakthrough at the time. The development of an elastic polyurethane continued as a substitute for rubber and was applied as aircrafts coating during World War II (Seymour & Kauffman, 1992). However, the real history of flexible polyurethane foams (PUF) began in 1954 with the first commercial production in Europe and soon in the United States (ISOPA, 2015). The success of commercial applications of PUF has established a market and has been growing ever since.

The properties of PUF depends on the formulation as it covers an extremely wide spectrum of stiffness, hardness, and densities (Gama, et al., 2018). Thus, segmenting the PUF market into flexible and rigid types. Rigid foams can be used as insulation panels, electronic components and sports equipment while flexible foams are used in upholstery, bedding and seat cushion. The increasing demand of PUF is a crucial PUF market driver in the coming years (Business Wire, 2019). It was estimated that the global PUF market would reach the valuation of USD 76,065.74 million during forecast period (2018-2023) and is expected to register a CAGR of 8.36% (Market Research Future, 2019).

The production of PUF requires a few basic components such as polyol(s), polyisocyanate(s), blowing agent and catalyst (Reymore, 1975). The chemical reaction involves two steps: gelling and blowing. In gelling, polyols react with polyisocyanates to produce urethane linkages. This is one of the contributions to the hard segment in the polymer network, where the polyol forms the soft segments (Shufen, et al., 2006). Hence, increased amount of isocyanate generates more harder segments, making the PUF more rigid. In blowing, water as a chemical blowing agent, reacts with polyisocyanates, producing urea which also contributes to the hard segment in the polymer network (Chattopadhyay, et al., 2006). In the same reaction, CO_2 gas is generated, and it acts as the major contributor in foaming the polyurethane. Although there are variety of blowing agents available, water is commonly used as it is abundant, non-toxic, cheap, and easier to implement in large scale production. Physical blowing agents are also utilised in industry. They expand the polymer by vaporisation of low boiling point compound, such as pentane and hexane (Gama, et al., 2018).

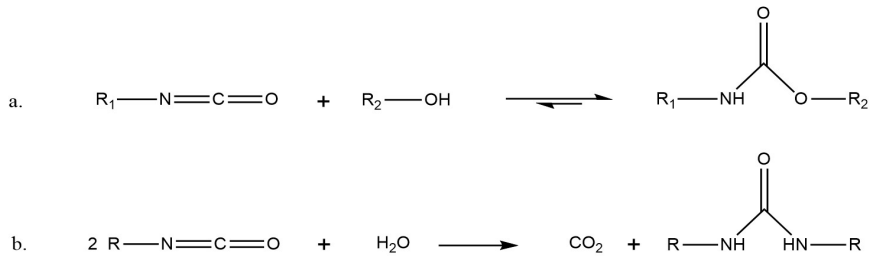


Figure 1 (a) Gelling reaction: isocyanate group reacts with hydroxyl group to form urethane group and **(b)** Blowing reaction: isocyanate groups react with water to form CO2 and urea group

MDI-based foams are typically more rigid than TDI-based foams as there exist a larger percentage of, the hard segment contributor, urea groups (Minnikanti, et al., 2011). Generally, increased molecular weight of polyols increase the flexibility of the foam, yielding higher tensile strength and elongation (Ashida, 2006). This suggests that high molecular weight polyol constructs the soft segment of polyurethane. The preparation of flexible and rigid foams involves similar chemistry, but the types of the polyols and isocyanates used will give varied characteristics of the foams as mentioned above.

The role of additives such as surfactants, catalysts, and chain extenders are also crucial in developing PUF. Surfactants are normally added in small proportions. It is used to lower surface tension, to improve the mixing of incompatible ingredients and to stabilise cell size (Lee & Ramesh, 2004). An example of surfactant that are commonly used is copolymer of polysiloxanes and alkylene oxide adducts. Selection of catalyst is important in order to enhance either the gelling or blowing reaction. The most common catalysts are amines and tin catalysts (Król, 2008). Finally, chain extenders can be used to improve resiliency, heat and chemical resistance properties by increasing the number of hard segments which are formed from isocyanate and chain extender (Blackwell, et al., 1981).

Applications of PUF in sportswear industry are limited to (1) shin pads, commonly used in contact sports including rugby, football and lacrosse, and (2) helmets worn by cyclist and skier to provide shock absorption and facilitate energy dissipation, thereby, protect players from serious injuries (Subic, 2019). However, these protective elements are said to be unfashionable and uncomfortable in the eyes of consumers, whether it is because of the bulkiness, heaviness or sweaty feeling that comes with using them. Nevertheless, players still wear them as they are required to comply with the rules of the competitive sport. To overcome such conflicts, the idea is to incorporate the protective elements into fabric to reduce the necessity of wearing external protective elements. A silicone-based polymer (SP) that has been incorporated into fabric was developed as an alternative to external protective elements. However, problem arose from SP as it is relatively heavy, and this is seen as a complication in sportswear due to sagging issues.

The aim of this study is to assess the applicability of PUF to replace SP that has been incorporated into fabrics. Aimed at sportswear market, PUF should be (1) lightweight, (2) can be moulded, (3) provide enough protection to withstand impact in sports, and (4) have the aesthetics to be competitive in the luxury fashion market. These are the objectives set up and the common goal for all formulations in this study.

Materials

Synthesis of polyurethane foam was carried out with the following ingredients: (1) polyols: 1,4-butanediol (BDO) (99%), 1,5-pentanediol (PDO) (96%), polyethylene glycol with molecular weights of 600 $g{\cdot}mol^{-1}$ and 8000 $g{\cdot}mol^{-1}$ (PEG 600 and PEG 8000); (2) diisocyanates: toluylene diisocyanate (TDI) (mixture of toluylene-2,6-diisocyanate and toluylene-2,4-diisocyanate isomers) (98%), 4,4'-methylenebis(phenyl isocyanate) (MDI) (98%); (3) catalyst: bis[2-(*N*,*N*-dimethylamino)ethyl] ether and (4) water.

Methods

Formation and moulding of PUF

Materials were prepared separately in two parts, Part A and Part B, before mixing. Part A consists of polyol(s), catalyst and water, whereas Part B consists of diisocyanate(s). In the formation step, after pouring in Part A, the solution was stirred continuously using a glass rod while pouring in Part B to ensure thorough mixing. The mixture was stirred for approximately 5-10 seconds just before it foamed and was held in a glass jar. For moulding, the mixture was quickly poured into a plastic/acrylic plastic mould until it occupied almost half of the void, keeping a margin for volume increase during foaming. All reactions were carried out at room temperature, except that when adding PEG 8000 as part of the formulation. Curing of the resulting PUF took place for 24 hours in room temperature. The procedure of preparing and mixing were carried out in the fume cupboard as the chemicals used were toxic and hazardous.

Note that PEG 8000 existed as solid at room temperature and it was placed in water bath at 80°C to be melted. The mixing of Part A and B was done relatively quickly before the mixture cools down, where PEG 8000 will solidify again. Thus, the reaction will take place rather spontaneously at this higher temperature of mixture.

Formulation

Different combinations of polyols and diisocyanates were used in preparing the PUF. Short chain alkanediols (BDO and PDO) were either used as the base or as chain extenders with polyethylene glycols (PEG 600 and PEG 8000) as the base polyols. When applied as base polyols, BDO and PDO were either used individually or in a mixture of 50 parts of BDO to 50 parts of PDO. In the case where PEG 600 and PEG 8000 were used as base polyols, they were used either individually or mixed from 400 to 1800 parts of PEG 600 to 100 parts of PEG 8000. In this study, 100 to 350 parts of diisocyanates, comprising of either TDI and MDI individually or a mixture of 50 parts TDI to 50 parts MDI, were mixed with 100 parts of polyols. As chain extenders, BDO was added in amounts ranging from 1 to 90 parts by weight per 100 parts of polyol. The diamine catalyst was employed in amounts ranging from 0.2 to 0.6 part by weight based on 100 parts by weight of polyol. Water, as the blowing agent, was present from 1 to 10 parts by weight per 100 parts of polyol and diisocyanate.

Figure 2 Compression test performed with EZ50

Performance testing of foam

Stress test was carried out with EZ50 as shown in Figure 2. During the experiments, the samples were analysed in compression mode to find out the amount of impact they can withstand. The maximum load applied was set at 45 kN with an extension rate of 2 mm·min^{-1}. The dimensions of the samples were measured and recorded before and after the experiments. The data were exported and processed using Excel.

Thermal potential of samples was evaluated by carrying out a straightforward test. Samples were put into oven at temperatures 37.5°C, 40°C, 60°C and 80°C respectively for 1 hour. A new parameter, deformation indicator (D_I), was introduced to demonstrate the degree of deformation of samples, solely through qualitative observations.

Results and Discussions

Iteration of formulations

Formulations were recorded and classified to three generations according to the parameters varied: types of polyols (PEG 600 or BDO) for the first generation; TDI mole ratio and weight percentage of chain extenders (i.e. BDO) for the second generation; and BDO wt% with PEG 8000 mole ratio for the third generation. Some formulations were not included as they were

trials conducted with PDO and/or MDI which produced PUFs that were too rigid or with properties that did not vary much with using BDO and/or TDI. Consequently, these formulations were not explored further in this study.

First generation experiments had utilised distinct types of polyol while varying TDI mole ratio. From Figure 3, mole ratio indicated the amount of TDI used relative to 1 mol of polyol. For PEG 600, the initial synthesis was done with 1.5 mole ratio and a sticky watery foam was observed. This foam was not preferable as the bubble size depleted during curing, so the formulation needed to be refined. As literature suggests that an increase in TDI mole ratio would harden the foam by increasing the hard segment of the PUF (Shufen, et al., 2006), a series of synthesis with higher TDI mole ratio was then carried out. From 1.7 to 1.85 mole ratio, the resulting foam was less watery but flaky. Then, a further increase to 1.9 caused the PUF to be a preferable foam that was neither flaky nor watery. However, the PUF was still considered relatively soft and could be refined further to improve hardness and resiliency.

In contrast, BDO produced hard foam for the initial synthesis with a 1 mole ratio of TDI. The increase in TDI mole ratio from 1.05 to 1.1 also resulted in hard foam. Similar results were reported in literature which short chain polyols reduced the mobility of the polymer structure, producing a stiff PUF (Hager & Helma, 1997).

In second generation experiments, the strategy was to harden the foam further, producing a more flexible and resilient foam. Chain extender such as BDO was introduced to further improve the strength of PUF synthesised from the first generation. Literature suggests that chain extenders would increase the number of urethane linkage, thus, further toughen the foam (Blackwell, et al., 1981). From Figure 4, the y-axis represents the weight percentage of BDO used per 1g of polyol. A lower weight percentage of BDO and a lower TDI mole ratio produced similar results as in the first generation, that is, watery and flaky. For formulations at 10, 15, and 20 wt% BDO content with 1.9 TDI mole ratio, soft PUFs are produced. However, no testing was done to compare them with first generation soft PUF as they were nearly similar from observation and touch.

As shown in Figure 4, a further increase of BDO from 30 to 50 wt% with a fix TDI mole ratio at 1.9 produced watery and sticky foam. It was then realised that the amount of TDI needed to be increased. The increment of TDI mole ratio to 2.5 produced paste-like foam and a further increase to 3.5 mole ratio resulted in hard PUF. This result implies the transition from soft to hard foam is a paste-like foam. At a much higher BDO content, which is 50 wt% and 90 wt%, the PUF turned out to be a much rigid foam, similar to that of the short chain polyol formulations in the first generation. Therefore, it is concluded that the plan of yielding a flexible yet resilient foam through this method was infeasible and a different approach was required.

In third generation experiments, the approach was to soften the hard foam produced from second generation by introducing a higher molecular weight polyol, that is, PEG 8000. Literature suggests that a lower molecular weight polyol makes rigid foam and an intermediate molecular weight polyol produces a flexible foam (Blackwell, et al., 1981). Mixtures of PEG 600 and PEG 8000 were formulated in third generation experiments. The idea was to have a mixture of PEG with an average molecular weight ranging from 1000 to $2000 g \cdot mol^{-1}$, providing an intermediate molecular weight.

From Figure 5, preliminary study of 0.056 mole ratio of PEG 8000 resulted in a sticky bubbly foam initially. After the foam was cured, the feature of the foam became waxy-like. It should also be noted that the PUF produced from third generation were all waxy-like foams, which were soft and weak.

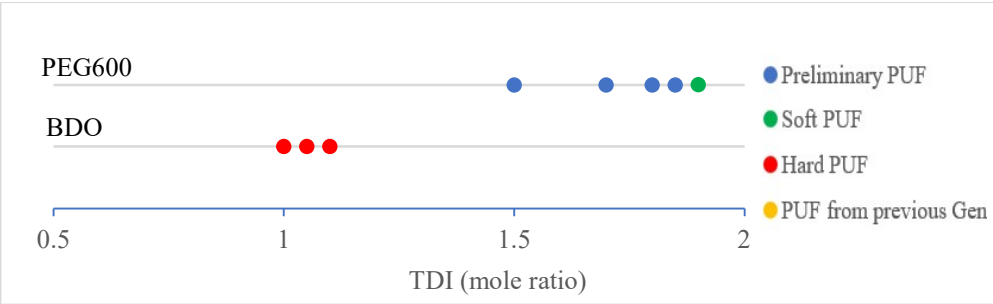


Figure 3 First generation formulation

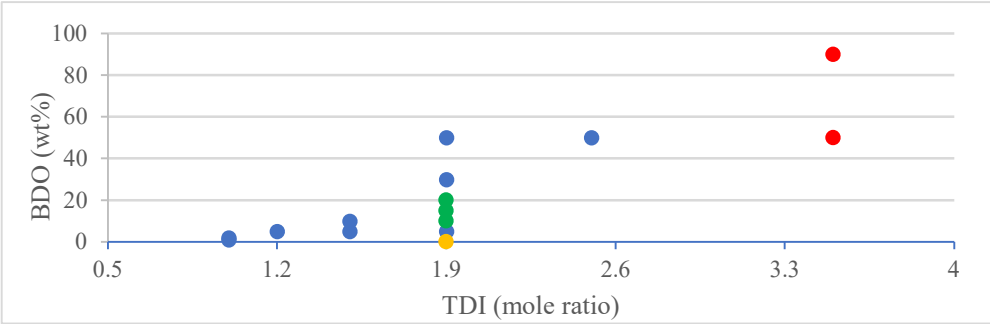


Figure 4 Second generation formulations

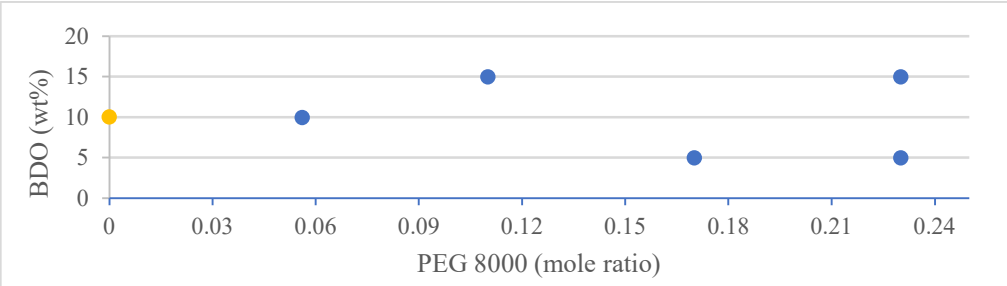


Figure 5 Third generation formulations

Mouldability and Aesthetic

To cater for the fashion industry, the developed PUF was expected to take the shape and details of a mould. Figure 6 shows a soft foam obtained from a formulation of PEG 600, TDI, BDO (1: 1.9: 30 wt%), whereas Figure 7 shows a rigid foam with the formulation of BDO, TDI (1: 1.05). From Figure 6 (a) and (b), it could be seen that the soft foam did not stay intact and fell apart when an attempt was made to take it out of the mould. The foam overflowed due to rise in volume and it is believed that greater void sizes made it difficult to take up the fine details of the mould. For the hard foam, it took the shape of a cylindrical glass jar, but it was brittle and broke when the shallower plastic mould was used. Thus, it is suggested that a deeper mould could be used to provide greater leeway for the increase of volume due to foaming of polyurethane, producing foam with higher consistency in shape.

Most of the PUF formed have irregular bubble size with rough surface, shown in Figure 8 (a) and (b). This may cause unpleasant appearance if it was considered as an aesthetic product. Distribution of bubbles and the regularity of bubble size were variables to be studied in the aesthetic component of the developed samples. In this case, orderly distributed bubbles and consistent bubble sizes were desired. Hence, it was suggested to add surfactants to the foam mixture to regulate the bubble sizes and to stabilise cell structure. This avoids the collapsing of bubbles and sub-surface voids (Randall, 2002), thus improving the presentation of the foam. Also, dye can be added into the mixture to provide colour and improve the external appearance of the foam.

Figure 6 Soft foam with formulation PEG 600, TDI, BDO (1: 1.9: 30 wt%) on a shallow plastic mould **(a)** before and **(b)** after an attempt of removing from mould.

Figure 7 Hard foam with formulation BDO, TDI (1: 1.05) after being removed from a deep mould (cylindrical glass jar).

Figure 8 (a) Closeup shot of foam surface, **(b)** Full shot of an acrylic mould covered in PUF

Formulation

Over 50 formulations were carried out and the proportion of each material was varied. Out of all the formulations, six series were chosen to be reviewed throughout the following discussion. PUFs were categorised based on their feature as either hard or soft. The compositions of individual raw materials for the PUFs chosen were listed in Table 1.

Table 1 Compositions of six chosen formulations

PUF	*Feature*	*BDO*	*PDO*	*PEG 600*	*PEG 8000*	*TDI*	*MDI*
A	*Hard*	*0*	*2*	*0*	*0*	*1.05*	*1.05*
B	*Hard*	*1*	*0*	*0*	*0*	*1.05*	*0*
C	*Hard*	*0*	*1*	*0*	*0*	*1.05*	*0*
D	*Soft*	*15 wt%**	*0*	*1*	*0*	*1.9*	*0*
E	*Soft*	*5 wt%**	*0*	*1*	*0.23*	*1.85*	*0*
F	*Soft*	*15 wt%**	*0*	*1*	*0.23*	*2.3*	*0*

Lightness

Synthesised foam has different densities depending on the proportions of Part A and Part B. Variation of densities for the distinct formulations is shown in Figure 9. From the figure, it could be seen that C has a lower density than B. This is explained by the longer chain alkanediols used in C, PDO instead of BDO, that resulted in more volume gain during the foaming process. It is considered as the rise in volume outweighed the increase in molecular weight, giving a lower density. Consequently, adding MDI will also increase the density of PUF. As shown in literature (Minnikanti, et al., 2011), the theoretical amount of urea in MDI based foams (35%) is significantly higher than TDI based foams (22%), hence resulting in a higher molecular weight of MDI-based foam, leading to greater density.

Figure 9 also shows the variation of soft PUFs' densities with higher molecular weight polyol and chain extenders. It was observed that adding higher molecular weight polyol i.e. PEG 8000 increased the density of foam compared to D. On the other hand, increasing chain extender content i.e. BDO from 5 wt% to 15 wt% also increased the density of foam. This is explained as greater amount of chain linkage was present in the polymer per unit volume, producing a denser PUF. Overall, all PUF formulations was lighter than silicone-based polymer that can prevent sagging issues in sportswear.

Mechanical Integrity

Literature showed that mechanical properties of PUF improved with increased density (Thirumal, et al., 2008). Compressive modulus of the materials were compared as it shows the stiffness of a solid material and often associated with its compressive strength.

Rigid PUFs (A, B and C) performed better than SP in general. When the density increased from 0.30 to 0.42 g·cm^{-3}, the compressive modulus of the PUFs also increased from 9.9 to 10.7 MPa. Similar results were reported for soft PUFs as shown in Figure 10. Minimum compressive modulus was demonstrated by D with the lowest density. Besides, the higher compressive modulus of soft PUFs (E and F) obtained from PEG 8000 could be attributed to the presence of chain extenders. Chain extenders are designed to react with isocyanate groups to become part of the hard segment. Therefore, increased BDO content from 5 wt% to 15 wt% produces additional hard segment and led to an increase in compressive modulus from 0.653 to 0.752 MPa.

The weakness of all of these PUF formulations was that they deformed after being compressed, not returning to their original shapes. Even though rigid PUFs were not as resilient as SP, they still perform better mechanically (higher compressive modulus) and thus providing better protection in sports.

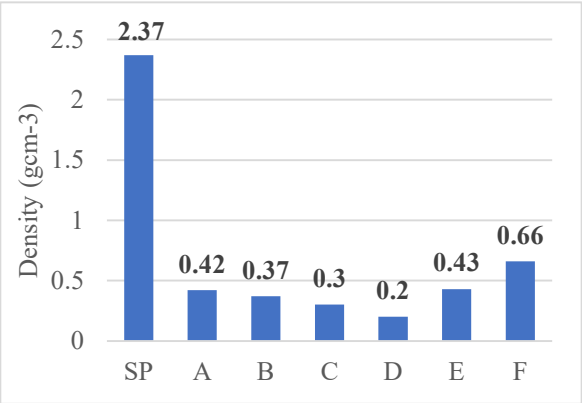


Figure 9 Density of SP and PUFs

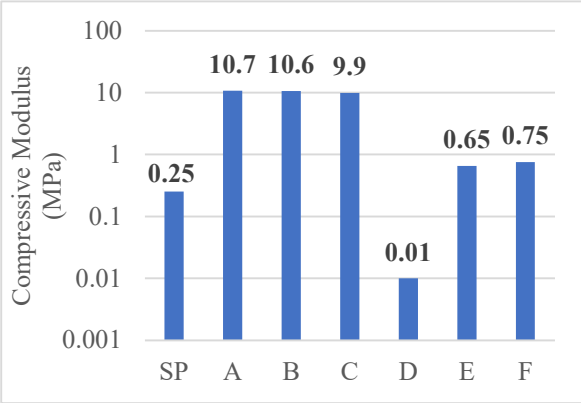


Figure 10 Compressive modulus of SP and PUFs

Thermal Stability

Deformation indicator (D_I) was introduced in the analysis of thermal performance. It is ranked from 1 to 10 based on qualitative judgements from observations, with 1 being no deformation at all (no change in mass and shape) and 10 being completely deformed (eg. melted). The PUF was assessed to simulate its thermal behaviour under normal body heat and elevated temperature environment i.e. washing machine. According to literature (Kébir, 2017), PUF is thermally more stable with increased repeating unit's molecular weight, that is, the molecular weight between urethane linkages. However, as shown in Figure 11, the thermal performance of soft PUFs (E and F) had notable effect at elevated temperature. The thermal stability was poor ($D_I = 10$), that is, they melted at 80°C when PEG 8000 was present. This result is explained with a scarce amount of hard segment (stiff urethane linkages and urea groups) and an abundance of the soft segment (high molecular weight polyol) present in the PUF where degradation of the soft segments had occurred before the hard segments. In this case, it can be explained that soft PUFs have a poorer thermal performance than rigid PUFs. Hence, it could be foreseen that rigid PUFs will perform better in the same range of temperature.

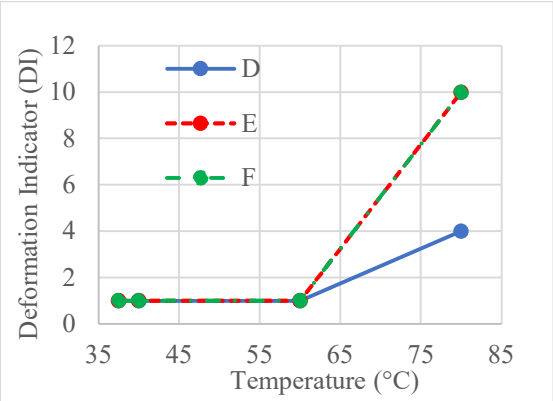


Figure 11 Deformation indicator (D_I) against temperature for different formulations

Conclusion

This preliminary study took a pragmatic approach in determining the correct formulation. The resulting samples were relatively sensitive to the proportions and types of polyols and diisocyanates used. From the developed samples, it can be concluded that PUF is a promising material as the substituent of SP mainly due to its lightweight property. This implies prevention of sagging complications in sportswear industry. Besides, rigid PUFs produced were stiffer than SP and hence protection from sports impact will be enhanced. Also, thermal performance of soft PUFs showed that normal body heat and elevated temperature environment have minimal impact on the feature of the foams, and hence a similar or better performance was predicted for rigid PUFs. However, the current formulations do not produce a PUF that has better resiliency than silicone-based polymer. Also, the formulations do not take up the shape and details of a mould, and its outward appearance need to be further enhanced to suit the fashion industry.

After numerous formulations using BDO, PDO, PEG 600 and PEG 8000 as polyols in the formulations, an intermediate molecular weight polyol can be applied to produce PUF with the strength and flexibility desired. For example, mixtures of different polyester polyol and polyether polyol with an average M_V of 2000 $gmol^{-1}$ can be further investigated. Besides, to increase the cross-link density and thus improving flexibility of PUFs, triols are worth to be explored in future work. Furthermore, surfactants such as different polysiloxanes and alkylene oxide adducts should also be studied to regulate and control the bubble size in order to obtain PUF with a consistent external appearance. Additionally, a characterisation of PUF can be carried out to give a better understanding of the structure and reaction kinetics. Examples of spectroscopy to be considered such as X-Ray Diffraction (XRD) and Fourier Transform Infrared (FTIR). In order to better simulate the impact in sports, ballistics test is suggested to be carried out with the developed PUF. Other tests such as water and air permeability test, water resistance test are also vital and is advised to be conducted if any mature PUF samples were to be launched in the sportswear market.

Acknowledgements

This study was inspired by Kiren Passi. The authors would like to offer their upmost gratitude to Kiren for her resources, in the form of fabrics, moulds and helpful discussions.

References

Ashida, K., 2006. *Polyurethane and Related Foams Chemistry and Technology.* Boca Raton, FL, USA: Taylor & Francis Group.

Blackwell, J., Nagarajan, M. R. & Hoitink, T. B., 1981. The Structure of the Hard Segments in MDI/diol/PTMA Polyurethane Elastomers. In: Washington, D.C: ACS Symposium Series. American Chemical Society, p. 179–196.

Business Wire, 2019. *Top 5 Vendors in the Global Polyurethane Foam Market 2019-2023.* [Online]
Available at: https://www.businesswire.com/news/home/20190711005310/en/Top-5-Vendors-Global-Polyurethane-Foam-Market
[Accessed 12 November 2019].

Chattopadhyay, D., Sreedhar, B. & K.V.S.N., R., 2006. Influence of varying hard segments on the properties of chemically crosslinked moisture-cured polyurethane-urea. In: s.l.:J. Polym. Sci. Part B Polym. Phys., pp. 102-118.

Gama, N., Ferreira, A. & Barros-Timmons, A., 2018. Polyurethane Foams: Past, Present, and Future.

Hager, S. L. & Helma, J. E. K. &. G. F., 1997. Polyether polyols suitable for flexible polyurethane foam prepared by co-initiation of aqueous solutions of solid polyhydroxyl initiators.

ISOPA, 2015. *The Benefits and history of polyurethane flexible foam.* [Online]
Available at: https://www.isopa.org/news/the-benefits-and-history-of-polyurethane-flexible-foam/
[Accessed 10 November 2019].

Kébir, N., 2017. Nonisocyanate thermoplastic polyurethane elastomers based on poly(ethylene glycol) prepared through the transurethanization approach. *Journal of Applied Polymer Science,* 134(45).

Koruyucu, A. & Durmaz, S., 2018. EXAMINATION OF SOME COMFORT RELEATED PROPERTIES OF SPORTSWEAR FABRIC COATED WITH POLYURETHANE FOAM. *International Journal of Latest Trends in Engineering and Technology,* 10(1), pp. 274-280.

Król, P., 2008. *Linear Polyurethanes: Synthesis Methods, Chemical Structures, Properties and Applications.* Leiden, The Netherlands: CRC Press.

Lee, S.-T. & Ramesh, N., 2004. *Polymeric Foams: Mechanisms and Materials.* New York: CRC Press.

Market Research Future, 2019. *Polyurethane Foam Market Trends, Industry Share, Growth Forecast, Business Strategy, Research Analysis on Competitive landscape and Key Vendors 2023.* [Online]
Available at: https://www.reuters.com/brandfeatures/venture-capital/article?id=92315
[Accessed 11 November 2019].

Minnikanti, V., Gamboa, R. & Motta, B., 2011. Low Density HR MDI Based Foams. Polyurethane Foam Association Meeting..

Mörck, R., Reimann, A. & Kringstad, K. P., 1989. Elastomeric Polyurethanes from a Kraft Lignin - Polyethylene Glycol, Toluene Diisocyanate System. Volume 30.

Randall, D. & Lee, S. eds., 2002. The Polyurethane Book. In: United Kingdom: Wiley, pp. 156-159.

Reymore, H. E. e. a., 1975. Isocyanu rate Foams: Chemistry, Properties and Processing. *Journal of Cellular Plastics,* p. 328–344.

Romanowski, P., n.d. Polyurethane. *How Products are Made,* Volume 6.

Seymour, R. B. & Kauffman, G. B., 1992. Polyurethanes: A class of modern versatile materials.

Shufen, L. et al., 2006. Studies on the thermal behavior of polyurethanes. *Polymer Plastic Technology and Engineering,* Volume 45, pp. 95-108.

Singh, S., 2002. *Blowing Agents for Polyurethane Foams.* Shawbury, UK: Rapra Technology Ltd.

Subic, A., 2019. *Materials in Sports Equipment.* s.l.:Woodhead Publishing.

Thirumal, M., Khastgir, D. & Singha, N. K., 2008. Effect of foam density on the properties of water blown rigid polyurethane foam. *Journal of Applied Polymer Science,* 108(3), pp. 1810-1817.

Tyler, D. & Venkatraman, P., 2012. Impact resistant materials and design principles for sportswear.

Wolska, A., Goździkiewicz, M. & Ryszkowska, J., 2012. Thermal and mechanical behaviour of flexible polyurethane foams modified with graphite and phosphorous fillers. *Journal of Materials Science,* 47(15), p. 5627.

A Numerical Investigation into the Role of Viscosity Contrast in the Convective Mixing of Carbon Dioxide and Oil for Geological Carbon Sequestration Applications

Kai L. Lim and Tricia W. T. Teo

Department of Chemical Engineering, Imperial College London, London SW7 2AZ, U.K.

Abstract

Carbon sequestration in deep geological formations such as depleted oil reservoirs forms a key step in the carbon capture and storage (CCS) process which is used to address climate change. The importance of convective mixing in this step makes it a topic of interest. This study aimed to numerically investigate the effect of temporal changes in the viscosity of supercritical carbon dioxide ($scCO_2$)-oil mixtures on the convective mixing of $scCO_2$ and oil. The model formulation involved developing an exponential viscosity-molar concentration correlation based on literature data. The Rayleigh number (Ra), a key parameter which measures the strength of convection to diffusion, was utilised. A two-dimensional (2D) Rayleigh-Bénard convection model was numerically implemented in COMSOL Multiphysics®. Three cases with an initial Ra value of 1000, and comprising a decreasing exponential, constant and increasing exponential viscosity-molar concentration profile respectively, were considered. A reference case for pure diffusion was also generated using MATLAB®. $ScCO_2$ concentration distribution plots and three mixing measures were used to evaluate the results. These showed that decreasing viscosity over time resulted in faster mixing and increased channelling while increasing viscosity had the opposite effect. These effects were similar to that of varying Ra with viscosity kept constant. An extended study performed found that beyond varying Ra, viscosity changes altered mixing patterns, and should thus be included in works involving the convective mixing of $scCO_2$ and oil.

Keywords: *COMSOL Multiphysics®, convective mixing, carbon capture and storage, convective fingering, Rayleigh number, dimensionless flux, dilution index, interfacial area, viscosity contrast, solubility trapping, supercritical CO_2*

Nomenclature

A_{geo}	Area of control volume geometry
A	Interfacial area
A_{diff}	Interfacial area for diffusive model
$\hat{A}$	Normalised interfacial area
c	Concentration (conc.) of solute (CO_2)
c_{top}	Conc. of saturated $scCO_2$-oil mixture
D	Diffusion coefficient
E	Dilution index
E_{max}	Maximum dilution index
F	Absolute flux
g	Gravitational constant
H	Characteristic height of reservoir
J	Dimensionless flux
κ	Permeability
M	Normalised dilution index
p	Pressure
ρ_{sol}	Density of dissolved CO_2 in oil
ρ_1	Density of saturated CO_2-oil mixture
ρ_0	Density of pure oil
$\Delta\rho_{max} = \rho_1 - \rho_0$	Maximum density difference
ϕ	Porosity
t	Absolute time
τ	Dimensionless time
T	Temperature
Ra	Rayleigh number
$Ra_{updated}$	Updated Rayleigh number
$\mathbf{u}$	Velocity vector
μ_{sol}	Viscosity of $scCO_2$-oil mixture
μ_0	Viscosity of pure oil
$\overline{\lvert\mu\rvert}$	Time- and space-weighted average viscosity
W	Width of reservoir
$\boldsymbol{x}(x, y)$	Position vector

1. Introduction

1.1 Context

The global energy requirement is predicted to increase by 80% from 2010 to 2050. This is expected to result in a 50% increase in greenhouse gas (GHG) emissions which contribute to global warming (OECD, 2012). Among the various GHGs, carbon dioxide (CO_2) has been identified as the largest contributor (Land Trust Alliance, 2018), and is projected to remain as a key driver of increases in global temperatures (IPCC, 2014).

Rising global temperatures have potentially far-reaching effects on human lives and the environment (Hayhoe *et al.,* 2017). This has prompted urgent calls-to-action from numerous organisations and global leaders to mitigate the consequences of increasing CO_2 emissions. Of these, the most notable is the Paris Agreement, which was formulated in 2016 and aims to limit global temperature rise in this century to less than 2°C above pre-industrial levels (UNCC, 2019). However, most strategies for boosting energy efficiency and reducing resource usage have been deemed insufficient to achieve this target.

New techniques of negative emission technologies present a glimmer of hope in the bid against global warming (Smith *et al.*, 2016). These include carbon capture and storage (CCS), which demonstrates the potential to capture up to 90% of CO_2 emissions (CCSA, 2016). This technology involves capturing CO_2 from its emission source, transporting it and storing it underground in deep geological formations such as depleted oil and gas fields, deep saline formations and coal mines.

1.2 Concepts

The pressure and temperature in the Earth's subsurface lie above the critical point of CO_2. As such, following its injection into the subsurface, CO_2 exists as a supercritical fluid which displaces the geological formation fluid.

However, being less dense and viscous than the geological formation fluid, supercritical CO_2 ($scCO_2$) eventually moves upwards until it hits an impermeable caprock at the formation top. This inhibits its escape in a process known as structural trapping (Javaheri *et al.*, 2009; Garcia, 2001; Gershenzon *et al.*, 2015).

However, fractures in the caprock pose a risk to $scCO_2$ leakage. The solubility trapping of $scCO_2$, which is described by the immobilisation of $scCO_2$ upon its dissolution in geological formation fluids (Newell *et al.*, 2018), mitigates this risk. Therefore, the dissolution of $scCO_2$ in these fluids is crucial for the permanent and secure storage of large quantities of $scCO_2$.

Two key mechanisms govern the dissolution process – molecular diffusion and density-driven convection. Using depleted oil reservoirs as an example, initially, at lower concentrations of $scCO_2$ in oil, the process is governed by molecular diffusion (Javaheri *et al.*, 2009). However, as the $scCO_2$ concentration in oil increases, the mixture's density increases. This creates a Rayleigh-Taylor instability in the system (Nazari Moghaddam *et al.*, 2012; Calabrese, 2017). At this point, density-driven convection occurs. Density differences between the $scCO_2$-oil mixture and the bulk pure oil induce convective currents that drive the mixture downwards. The rate at which further growth of perturbations occurs is moderated by viscosity (Gallaire and Brun, 2017). Over time, complex finger patterns known as convective fingerings form, which accelerate the dissolution process.

Between diffusion and density-driven convection, the latter involves a larger portion of the total $scCO_2$ quantity (Nazari *et al.*, 2015) and expedites the dissolution process (Liyanage *et al.*, 2018). It is also a crucial transport mechanism that brings $scCO_2$ deeper into the formation, thereby increasing the area covered by it (Simmons *et al.*, 2001). Thus, density-driven convection is a key mechanism that provides a faster, more stable and long-term storage solution (Wang *et al.*, 2017), and it is worthwhile to study the convective mixing process in the context of CO_2 solubility trapping.

1.3 Literature review and objectives

Various works have attempted to study the convective mixing process numerically.

Most numerical experiments in the current literature utilise the Rayleigh-Taylor model which models a closed two-phase system (Neufeld *et al.*, 2010; Liyanage, 2018). However, based on numerical simulations performed by Lindeberg and Bergmo (2002) on $scCO_2$ and brine, the interface between $scCO_2$-rich brine and pure brine was found to be unaffected by the formation of fingers. This indicated the presence of a local equilibrium between $scCO_2$-rich brine and pure brine. Hence, a simplification can be applied: an open one-phase system, comprising a sharp and rigid interface with constant concentration at the top boundary, can be used instead, corresponding to a Rayleigh-Bénard model (Ghesmat *et al.*, 2011).

A range of studies implemented a two-dimensional (2D) Rayleigh-Bénard model to study convective mixing. Notable examples include a study conducted by Hidalgo and Carrera (2009) on how dispersion affected the onset time of convection and another by Pau *et al.*, (2010) which discovered that the constant flux regime is predominated by the convection-dominated flow regime.

A key parameter which is used in studies on convective mixing is the Rayleigh number (Ra) (Emami-Meybodi *et al.*, 2015; Liyanage, 2018). This measures the strength of convection to diffusion and is defined below.

$$Ra = \frac{\kappa \Delta \rho g H}{\mu \phi D} \tag{1}$$

Changes in density and viscosity affect the convective mixing process as explained in Section 1.2. Numerous works have investigated the convective mixing of $scCO_2$-brine mixtures with a focus on changes in density while viscosity was kept constant (Emami-Meybodi *et al.*, 2015; Riaz and Cinar, 2014). These were mainly for applications in deep saline aquifers. However, there exists a knowledge gap in the convective mixing behaviour of $scCO_2$-oil mixtures given a changing viscosity profile over time. As such, for applications in depleted oil reservoirs, this study aims to numerically investigate how temporal changes in the viscosity of $scCO_2$-oil mixtures, due to increasing $scCO_2$ concentration, affects convective mixing.

This paper is laid out as follows: Section 2 details the model formulation, numerical implementation and data processing methods used to study the problem. The numerical results are presented and discussed in Section 3. In Section 4, an extended study based on the findings from Section 3 is discussed. Finally, Section 5 concludes the study and highlights areas for future work.

2. Theory and Methodology

2.1 Fluid properties

A key step in the model formulation involved determining an empirical correlation for changes in the mixture's (i) density with $scCO_2$ concentration and (ii) viscosity with $scCO_2$ concentration. To this end, a literature review to obtain fluid property data was conducted. Regression analysis involving linear and non-linear least-squares (R^2) fitting was employed.

2.1.1 Density-concentration correlation

Data points from three data sets were found (Calabrese, 2017; Ciotta *et al.*, 2009) and normalised using $\frac{\rho_{sol}-\rho_0}{\Delta\rho_{max}}$ (refer to Appendix B for details on normalisation). A linear and a polynomial correlation were obtained as depicted in Figure 1. The linear correlation shown in Equation (2) was chosen despite its lower R^2 value for ease of comparison with previous studies done by Eckel (2019) on scCO2-brine mixtures.

$$\frac{\rho_{sol}-\rho_0}{\Delta\rho_{max}} = 4 \times 10^{-4} c \tag{2}$$

The terms were re-arranged to obtain the equation below.

$$\rho_{sol} = \Delta\rho_{max}(4 \times 10^{-4} c) + \rho_0 \tag{3}$$

2.1.2 Viscosity-concentration correlation

The data points from nine data sets (Calabrese, 2017; Ciotta *et al.*, 2009; Yongmao *et al.*, 2004) were normalised with $\frac{\mu}{\mu_o}$ (refer to Appendix B for details on normalisation) as shown in Figure 2. The decreasing exponential correlation, expressed in the equation below, was chosen due to its higher R^2 value.

$$\frac{\mu_{sol}}{\mu_0} = exp\,(-4 \times 10^{-4} c) \tag{4}$$

Again, the terms were re-arranged as shown in the equation below.

$$\mu_{sol} = \mu_0 exp\,(-4 \times 10^{-4} c) \quad (5)$$

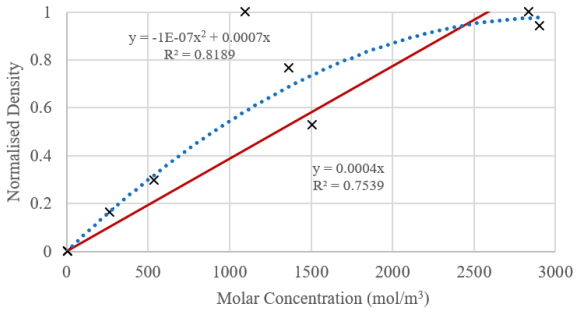


Figure 1. *Normalised density against molar concentration*

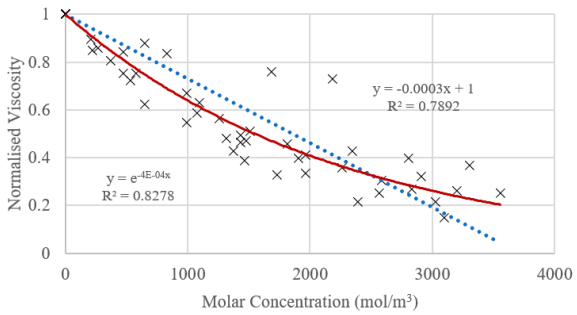


Figure 2. *Normalised viscosity against molar concentration*

2.2 Model Parameter Values

Ra, the key parameter introduced in Section 1.3, was employed in this study. A literature review on depleted oil fields found globally was performed to obtain values for the various parameters in Equation (1). As shown in Figure 3, most depleted oil reservoirs had Ra values in the range of 1000 to 10000 (refer to Appendices A and C for the details). A Ra value of 1000 was chosen for this study due to time limitations. The list of literature values implemented in the models are shown in Table 1.

Table 1. *Literature values to be used for the models*

Parameter	Values
c_{top} $(mol\,m^{-3})$	2500
p_0 (atm)	237
T (K)	358
κ (m^2)	2.92×10^{-14}
ϕ	0.216
D (s^{-1})	4.83×10^{-9}
$\Delta\rho_{max}$ $(kg\,m^{-3})$	16.2
μ_0 $(Pa\,s)$	1.74×10^{-3}
H (m)	400
g $(m\,s^{-2})$	9.81
Ra	1000

2.3 Numerical model for convective mixing

2.3.1 Governing equations and assumptions

The convective mixing problem can be formulated mathematically with the following three governing equations.

Continuity equation: $$\phi \frac{\partial}{\partial t}\rho + \nabla \cdot (\rho \boldsymbol{u}) = 0 \quad (6)$$

Darcy's law: $$\boldsymbol{u} = -\frac{k}{\mu_{sol}}(\boldsymbol{\nabla p} + \rho \boldsymbol{g}),\ \boldsymbol{g} = \begin{bmatrix} 9.81 \\ 0 \end{bmatrix} \quad (7)$$

Solute transport equation: $$\phi \frac{\partial}{\partial t} c + \nabla \cdot (-\phi D \nabla c + \boldsymbol{u} c) = 0 \quad (8)$$

Equation (6) describes mass and momentum conservation while Equations (7) and (8) describe the fluid flow and diffusive flux respectively. These, together with Equations (3) and (5) ensure a well-posed problem.

Drawing upon literature, a 2D Rayleigh-Bénard convection model for a homogenous porous medium was used to describe the problem. The following assumptions were made:

- <u>Fluid flow</u>

- Incompressible system with laminar flow regime[1]

- <u>Solute transport</u>

- Molecular diffusion described using Fick's second law
- Constant equilibrium concentration of saturated $scCO_2$-oil mixture at mixing interface in a one-phase system
- Negligible velocity-induced dispersion effects, geochemistry effects and molecular interactions

- <u>Porous media properties</u>

- Constant and isotropic ϕ and κ
- Solvent fully saturates the media
- Tortuosity assumed as unity

2.3.2 Initial (ICs) and boundary conditions (BCs)

Figure 4 shows the control volume that represents the problem of interest, with the imposed ICs and BCs. At IC $t = 0$, the bulk fluid, pure oil, is motionless. Therefore, $c = 0$ and $\boldsymbol{u} = 0$. The pressure at any point is described using the model-adapted hydrostatic equation below.

Hydrostatic equation: $$p = \rho_{sol} g\,(H - y) + p_0 \quad (9)$$

A reference pressure, p_0, was defined at point (0,400) to obtain a unique solution.

As shown in Figure 4, the no-flow BC was imposed on all boundaries while the no-flux BC was imposed on the left, bottom and right boundaries. Conforming to the Rayleigh-Bénard model, a constant concentration Dirichlet BC, c_0, was placed on the top boundary. This concentration corresponds to that of a saturated $scCO_2$-oil mixture.

2.3.3 Numerical implementation

The 2D Rayleigh-Bénard convection model was implemented in COMSOL Multiphysics® (Version 5.3), a finite element simulation software with computational fluid dynamics (CFD) technology. This software was selected as it provides the flexibility to modify governing equations or introduce partial differential equations (PDEs) while maintaining a user-friendly interface and a wide range of online documentation. Furthermore, advancements in CFD technology have made CFD a key tool for accurate and relatively inexpensive analysis of fluid flow problems (Zhang and Agarwal, 2014; BMT Fluid Mechanics, 2010).

[1] The laminar flow assumption was verified post-simulation as calculations gave $Re = \frac{\rho|u_y|d(1-\phi)}{\mu} < 10$ for $Ra = 1000$. The grain size (d) was approximated using the Carman-Kozeny equation (Materazzi, 2017).

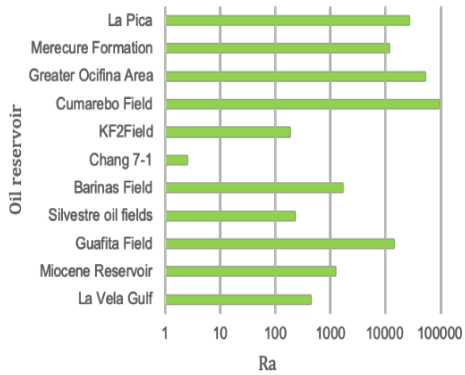


Figure 3. *Ra for various oil fields*

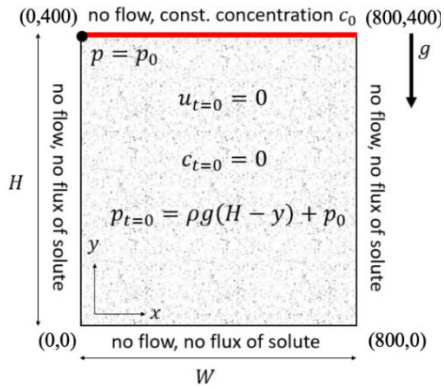


Figure 4. *Control volume for convective mixing of CO_2 in oil reservoir*

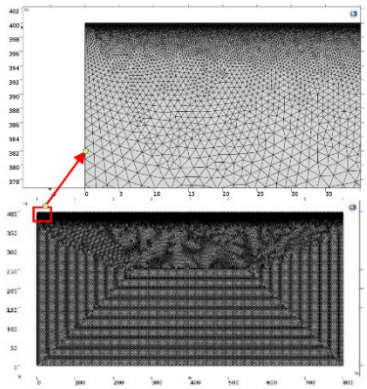


Figure 5. *Left: section of mesh, right: geometry used for control volume*

The 'Darcy's Law' physics interface from the 'Porous Media and Subsurface Flow' module was employed to relate the pressure gradient and flow rate in the reservoir while the 'Transport of Diluted Species' interface from the 'Chemical Species Transport' module was utilised to analyse the concentration of $scCO_2$ in oil. A two-way coupled problem was formulated using both interfaces.

The mesh for the control volume was created using triangular elements as they offered the ease of varying resolutions within the area using an automatic mesh generator (Perelmuter and Slivker, 2003). A study on the Grid Convergence Index (GCI), which estimates how changes in mesh size would affect the results, showed that the "finest" settings gave errors below 5.5% (Eckel, 2019). This implied that further refinements of the mesh would require much longer computational times for small improvements in results.

The mesh was further refined only near the top boundary where fingers were expected to form. This gave a resolution of 0.1m at the top boundary and 4m at the bottom boundary (refer to Figure 5). In order to induce an instability, a perturbation is required at the top boundary. Methods for this include creating a concentration disturbance (Newell et al., 2018). In this study, employing triangular elements that increase in size from the top to bottom boundary created an intrinsic perturbation.

Referring to Figure 4, the height of the control volume was set as 400m to simulate the average thickness of depleted oil reservoirs. As the width was not used in calculating Ra, its value was not crucial and was thus chosen to be 800m simply to ensure sufficient length to observe the mixing process.

The simulations were solved using the MUltifrontal Massively Parallel Sparse (MUMPS) direct solver, which employs Lower-Upper (LU) decomposition. The Backward Differentiation Formula (BDF) solver was used for time-stepping due to its stability (COMSOL Blog, 2013). An adaptive time stepping method was adopted, which automatically modified the time step to keep within the desired tolerance. The simulations each took between 30- and 50-hours wall clock time and were stopped only after convective shutdown was achieved. Their time scales are shown in Table 2.

2.4 Analytical solution for pure diffusion

As a reference for the difference between the performance of a pure diffusion and a convection-diffusion case, the PDE for pure diffusion shown below was solved analytically in MATLAB®. This result was compared with the solutions obtained from COMSOL®.

1D diffusion in porous media: $$\frac{\partial c}{\partial t} = D\nabla^2 c \qquad (10)$$

2.5 Data processing

The numerical results were analysed in COMSOL® using the metrics in the following sections.

2.5.1 Concentration distribution plots and profiles

2D surface plots depicting $\frac{c}{c_{top}}$ at specific times were generated in COMSOL® for visualisation of the CO_2 concentration distribution within the control volume. The plots' scale ranged from dark blue to red corresponding to $\frac{c}{c_{top}}$ of 0 to 1. Frequency distribution profiles, which correspond to the surface plots, were also created for a more quantitative analysis. The frequencies were normalised by the area of the 2D system (A_{geo}) to give a scale of 0 to 1.

2.5.2 Dimensionless flux

A key global measure employed to quantify the degree of mixing was the dissolution flux per unit width. This quantifies the rate of change of total dissolved $scCO_2$ in oil and was computed in COMSOL® using the following equation:

$$F(t) = \frac{1}{W}\frac{d}{dt}\int_{A(t)} c(W, (H-y), t)dW \qquad (11)$$

For comparisons across simulations, $J(\tau)$, the dimensionless form of the dissolution flux and a corresponding dimensionless time τ were used as shown below.

$$J(\tau) = \frac{H}{D\phi c^*}F \qquad (12)$$

$$\tau = \frac{Dt}{H^2} \qquad (13)$$

2.5.3 Normalised dilution index

Another approach to evaluate mixing performance entails the use of the dilution index (E) proposed by Kitanidis (1994). This measures the degree of uniformity of solute mass distribution within the system through a combined evaluation of the solute volume and the homogeneity of the concentration with the equation:

$$E(t) = exp[-\int_{A_{geo}} p(\boldsymbol{x}, t)\, ln\big(p(\boldsymbol{x}, t)\big)\, dA_{geo}] \quad (14)$$

where $p(\boldsymbol{x}, t)$ refers to the probability distribution of the position of a solute particle at time t. The continuous form of the equation was used as the mesh employed for the simulations was sufficiently fine to define a continuous system which could be evaluated using COMSOL®.

The computed E values were normalised against the maximum value (E_{max}), which coincided A_{geo}. This gave a scale of 0 to 1, which represents minimum to maximum dilution.

The normalised dilution index (M) was obtained using the equation below.

$$M = \frac{E}{E_{max}} = \frac{E}{A_{geo}} \quad (15)$$

2.5.4 Normalised interfacial area

As explained in Section 1.2, convection drives $scCO_2$-rich oil downwards, forming convective fingerings, which increase the interfacial area between $scCO_2$ and oil. The interfacial area, A, is shown by the green contours in Figure 6, and can be employed as a measure to determine the degree of mixing due to convection. The larger the interfacial area, the higher the rate of mass transfer (Kitanidis, 1994). An integration function in COMSOL® was utilised to compute A across all time and at a chosen concentration of $\frac{c}{c_{top}} = 0.1$. This was then normalised by A_{diff}, which corresponded to the width of the control volume, as shown in the equation below.

$$\hat{A} = \frac{A}{A_{diff}} \quad (16)$$

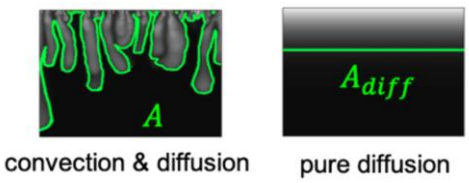


Figure 6. *Left: interfacial area for a convective model, right: interfacial area for a purely diffusive model*

2.6 Case studies

Four cases were considered and numerically implemented in COMSOL® to investigate how changes in the viscosity of $scCO_2$-oil mixtures over time affect convective mixing. A summary of the cases is presented in Table 2. As a reference (see Section 2.4), a diffusive model ($Model_{diff}$) with a Ra value of 1000 was generated in MATLAB®. $Model_{dec}$ corresponded to the $scCO_2$-oil mixing problem of interest. Its viscosity profile was obtained from Section 2.1.2. $Model_{con}$ provided a base case for comparison while $Model_{inc}$ was generated to draw comparisons to studies done on $scCO_2$-brine convective mixing where viscosity increased with concentration. In these three cases, μ_0 was used as the initial viscosity. $Model_{ext}$ was created as part of an extended study conducted following the analysis of the results of $Model_{dec}$ to $Model_{inc}$, and will be elaborated upon in Section 4.

Table 2. *Summary of cases studied*

Case	Temporal viscosity variation	Ra	Viscosity profile	Time scale/ $\times 10^3$ (a)
$Model_{dec}$	Decreasing	1000	$\mu = \mu_0 e^{-4\times10^{-4}c}$	75
$Model_{con}$	Constant	1000	$\mu = \mu_0$	75
$Model_{inc}$	Increasing	1000	$\mu = 2\mu_0 - \mu_0 e^{-4\times10^{-4}c}$	100
$Model_{ext}$	Constant	2000	$\mu = \overline{\lvert\mu\rvert}$	50

3. Results and Discussion

3.1 Concentration distribution plots and profiles

3.1.1 Overview of convective regimes

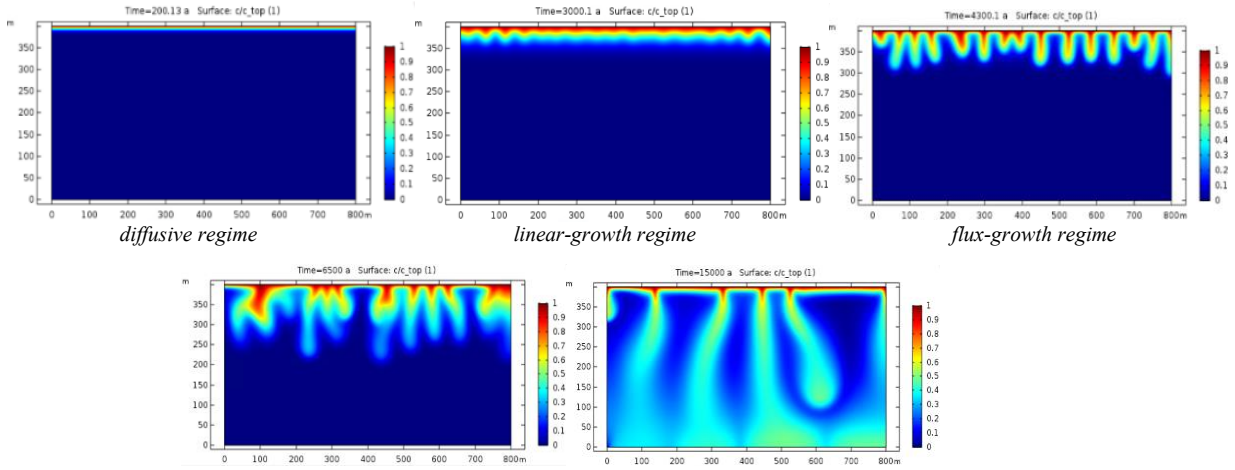
Across $Model_{dec}$ to $Model_{inc}$, the dissolution of $scCO_2$ in oil was observed to adhere to five distinct convective regimes albeit at different rates. These regimes were clearly distinguishable from the 2D concentration distribution surface plots generated from COMSOL®. Referring to Figure 7, in the initial diffusive regime, a high concentration gradient existed at the interface, and the $scCO_2$-oil mixture moved down as a horizontal front by molecular diffusion. As the diffusive boundary grows, instability was initiated. The system moved into the linear-growth regime and the front became wavy. After a critical time, the perturbations in the system set off convection, and the flux-growth regime was observed. Convective fingers appeared, disrupting the diffusive boundary layer, and the dissolution flux began to increase until a maximum flux was reached. After this point, the fingers began to merge and thicken, resulting in complex downwelling plumes. These corresponded to the constant flux regime. Eventually, the finger plumes reached the bottom boundary and interactions between the upwelling and downwelling plumes led to saturation of the system. These constituted the convective shutdown regime.

3.1.2 Comparison at constant τ

2D surface plots at constant selected times across the time scales of the three cases were created and compared. An overview of the comparisons at constant τ can be found in Appendix D.

The results for $\tau = 6.38\, x\, 10^{-3}$ will be analysed in detail here. This corresponds to the time when the finger plumes of $Model_{con}$ first reach a penetration depth of 200m. This depth represents the half-height of the control volume and would allow for the development of sufficient mixing features for analysis.

Referring to Table 3, at this value of τ, the 2D plots for all three cases showed finger structures, implying that all the systems had progressed past the diffusive and linear-growth regimes. From $Model_{dec}$ to $Model_{inc}$, the average penetration depths showed a decreasing trend, suggesting that the rate of mixing decreased with increasing viscosity contrast and vice versa. Specifically, with reference to Table 3, the convective shutdown regime where well-developed finger plumes spanned the full height of the control volume was observed in $Model_{dec}$. The 2D plot showed significant mixing with large regions of light blue and green within the fingers, and comparatively smaller red areas when compared to the other two cases. This



diffusive regime *linear-growth regime* *flux-growth regime*

constant flux regime *convective shutdown regime*

Figure 7. *Convective regimes observed for the dissolution of CO_2 in oil*

corresponded to Figure 8, where the curve for $Model_{dec}$ showed higher frequencies of $\frac{c}{c_{top}}$ between 0 and 0.5, and lower frequencies of $\frac{c}{c_{top}}$ between 0.6 and 1, as compared to the other cases.

In $Model_{con}$, the later stage of the constant flux regime was seen, and the fingers had an approximate length of 200m as depicted in Table 3. Faded light blue fingertips were observed, giving rise to higher frequencies of $\frac{c}{c_{top}}$ between 0.1 and 0.5 than $Model_{inc}$ and $Model_{diff}$ in Figure 8.

The flux-growth regime was observed in $Model_{inc}$, with fingers showing an approximate length of 100m in Table 3. Most of the oil in the control volume remained unmixed. As such, as seen in Figure 8, the frequency distribution profile for $Model_{inc}$ resembles that of $Model_{diff}$. The wider fingers seen in the 2D plot for $Model_{inc}$ should not be confused with merging – a phenomenon typically observed in the constant flux regime[2].

The observed trend for the rates of mixing of the three cases falls within expectations and will be discussed in Section 3.2.1.

3.1.3 Comparison at constant finger penetration depth

A comparison of the concentration distributions at selected finger penetration depths for the three cases was also performed, and an overview can be found in Appendix D. This comparison was used to determine whether the three cases showed variations in mixing behaviour when the differences in their mixing rates was discounted.

The results for a penetration depth of 200m will be examined in detail. This depth was chosen for consistency with Section 3.1.2.

As shown in Table 3, $Model_{dec}$ took approximately half the duration taken by $Model_{con}$ to reach a depth of 200m while $Model_{inc}$ took approximately 1.3 times, consistent with the findings from Section 3.1.2. At this depth, the constant flux regime was observed in all cases.

The three cases showed similar number of fingers at half-depth, as evident from the 2D plots in Table 3. This was interesting to note given that the number of fingers varied from $Model_{dec}$ to $Model_{inc}$ in the flux-growth regime. A possible explanation for this is as such: although $Model_{dec}$ had more fingers in the flux-growth regime, these fingers were closer to one another such that the bulk fluid posed less resistance to sideway interactions between neighbouring fingers, thereby promoting finger coalescence and reducing the total number of fingers in the constant flux regime. Conversely, $Model_{inc}$ had fewer fingers in the flux-growth regime but the fingers had to travel a longer distance sideway, making finger coalescence effortful and less likely (Zimmerman and Homsy, 1992). Therefore, the decrease in number of mergers between initial fingers resulted in the similar number of fingers at half-depth.

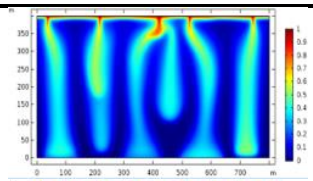
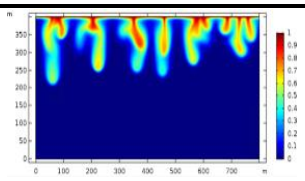
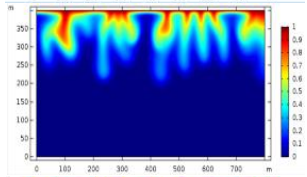
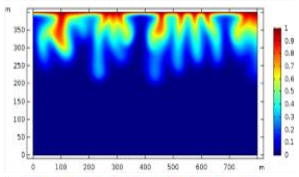
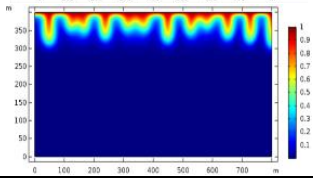
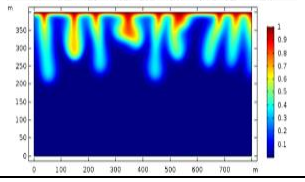
Another interesting observation to highlight is the definition of the finger contours. As depicted in Table 3, $Model_{dec}$ showed the most well-defined finger contours among the three cases. As such, there were fewer light blue regions around the fingers, which corresponded to lower frequencies of $\frac{c}{c_{top}}$ between 0.1 and 0.4 as plotted in Figure 9. It is likely that the well-defined contours developed because the more viscous bulk fluid posed resistance to the outward diffusion of the less viscous fingers.

In $Model_{con}$, faded fingertips with vague contours were identifiable. These corresponded to more light blue regions and thus higher frequencies of $\frac{c}{c_{top}}$ between 0.1 and 0.3 than $Model_{dec}$ as depicted in Figure 9. In this case, the fading phenomenon could be explained by the longer development time of $\tau = 6.38\text{x}\ 10^{-3}$. The absence of a resistive barrier due to viscosity contrast between the fingers and the bulk fluid enabled the diffusion of $scCO_2$ into the bulk fluid.

Interestingly, the fingers in the 2D plots of $Model_{inc}$ exhibited less fingertip fading than that of $Model_{con}$. This correlated to slightly higher frequencies of $\frac{c}{c_{top}}$ between 0.35 and 0.4 for $Model_{inc}$ in Figure 9. It is plausible that

[2] λ_c, the critical wavelength before the onset of convection, is proportional to the system's μ (Tang *et al.*, 2019). In case 3, due to a higher μ, λ_c is larger. Thus, the fingers in the 2D plot are wider, and appear to merge although the system is in the flux-growth regime.

Table 3. *Comparison of concentration distribution of the three cases at constant τ and constant penetration depth*

Case	$\frac{c}{c_{top}}$ at $\tau = 6.38\ x\ 10^{-3}$	$\frac{c}{c_{top}}$ at $200m$	τ/ x 10^{-3} at $200m$
1			3.24
2			6.38
3			8.46

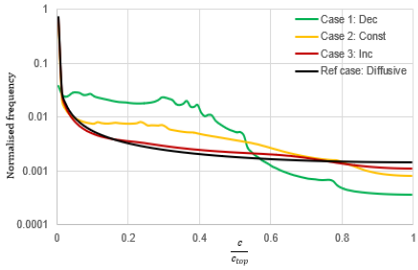


Figure 8. *Frequency distribution profiles at* $\tau = 6.38\ x\ 10^{-3}$

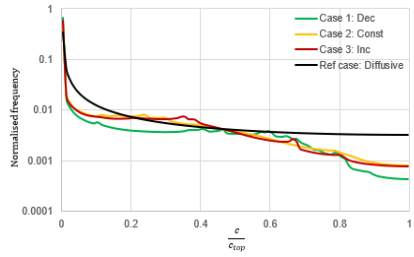


Figure 9. *Frequency distribution profiles at 200m*

due to the higher viscosity of the fingers in $Model_{inc}$, there was greater inertia for diffusion to occur, resulting in less faint contours than that observed in $Model_{con}$.

Overall, as evidenced from Figure 9, all three models showed lower frequencies of $\frac{c}{c_{top}}$ between 0 and 0.2, and higher frequencies of $\frac{c}{c_{top}}$ between 0.7 and 1 when compared to $Model_{diff}$. The three viscosity models also showed rather similar concentration distribution profiles at the same penetration depth, which implied approximately similar mixing performance among them. Thus, there did not seem to be a compromise on mixing performance due to development times; $Model_{dec}$ did not perform more poorly due to its quicker development time.

3.2 Mixing measures

3.2.1 Dimensionless flux

The graphs of J against τ for $Model_{dec}$ to $Model_{inc}$ and $Model_{diff}$ are shown in Figure 10. Five distinct regions are clearly identifiable from the graphs, corresponding to the five convective regimes mentioned in Section 3.1.1 (refer to Appendix E for details).

For applications in CO_2 storage, a high maximum dimensionless flux is desirable as this indicates a high rate of $scCO_2$ dissolution in oil. In addition, convection should commence as early as possible and hence it is crucial to work out the convection onset times. Points indicating these two parameters are shown with crosses and a triangle symbols respectively on Figure 10. It is evident that the earliest onset of convection and maximum flux occurred in $Model_{dec}$. The ratios between $Model_{dec}$ and $Model_{inc}$ with respect to $Model_{const}$ for these two parameters are shown in Table 4, with absolute values given in Appendix F. Similarities can be drawn to a study done by Emami-Maybodi *et al.* (2015) for $scCO_2$-brine mixtures where the viscosity was kept constant while Ra was varied (Figure 11). The trend observed for $Model_{dec}$ was comparable to an increase in Ra, and vice versa for $Model_{inc}$. This similarity was expected as Ra varies inversely with viscosity. Thus, a decrease in viscosity over time would result in an increase in Ra, leading to a stronger convection term which contributed to the higher $J(\tau)$ and earlier onset of convection for $Model_{dec}$. The similarity to Emami-Meybodi's (2015) study provided motivation for an extended study to be conducted to explore whether the effect of changing viscosity on convective mixing behaviours could be captured simply by changing Ra. This will be expanded on in Section 4.

Table 4. *Parameter comparisons for the various mixing measures*

Mixing measure	Parameter	$\frac{Model_{dec}}{Model_{const}}$	$\frac{Model_{inc}}{Model_{const}}$
$J(\tau)$	Max $J(\tau)$	1.645	0.770
	Time of convection onset / τ	0.352	1.825
M		0.365	1.737
$\hat{A}$	Max $\hat{A}$	1.392	0.791

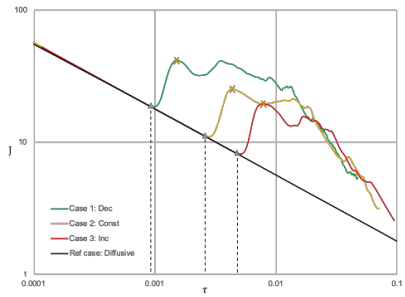


Figure 10. *Dimensionless flux against dimensionless time*

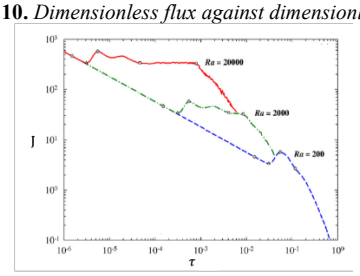


Figure 11. *Study done by Emami-Meybodi et al. (2015) showing dimensionless flux against dimensionless time for CO_2-brine mixtures*

3.2.2 Normalised dilution index

The normalised dilution index for all models were evaluated and plotted in Figure 12. The convection onset times are shown with a triangle symbol. At these points in time, the curves for $Model_{dec}$ to $Model_{inc}$ deviated from $Model_{diff}$. It is worthwhile to highlight that the convection onset times observed were similar to that in Section 3.2.1, as evidenced from Table 4 (refer to Appendix F for the absolute times). This observation was expected given that with the onset of convection, the dominant mode of mass transport shifts from diffusion to convection, resulting in an increase in the dissolution flux, and thus, the dilution index as the denser $scCO_2$-oil mixture is brought away from the interface.

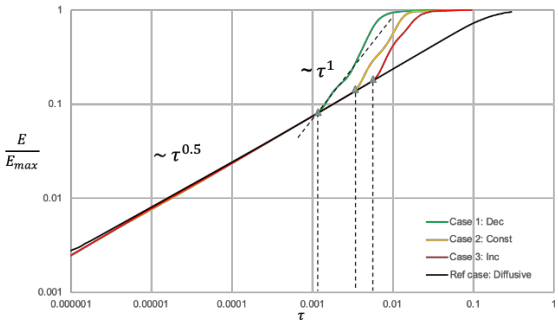


Figure 12. *Normalised dilution index against dimensionless time*

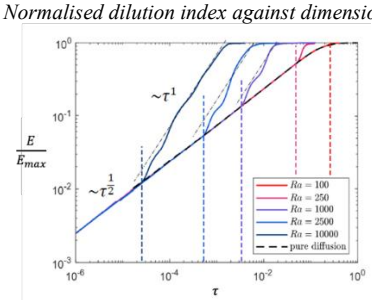


Figure 13. *Study done by Eckel (2019) showing normalised dilution index against dimensionless time for CO_2-brine mixtures*

In all three cases, the dilution index was observed to increase according to the following equation:

$$E \sim \tau^{n/2} \tag{17}$$

where n represents the dimension of the domain ($n = 1$ for pure diffusion and $n = 2$ for diffusion and convection). Hence, the dilution index increased with $\sim\tau^{0.5}$ for the diffusive case and $\sim\tau^1$ after the onset of convection. These gradients were consistent among all cases, which was expected as they correspond to the growth of a Gaussian plume over time (Kitanidis, 1994).

The trends noted above were similar to a study done by Eckel (2019) on $scCO_2$-brine mixtures as shown in Figure 13, where viscosity was maintained constant while Ra was varied. $Model_{dec}$ was observed to show comparable features to an increase in Ra while $Model_{inc}$ showed similarities to a reduction in Ra. These observations provided further motivation for an extended study to be performed to determine whether the effect of changing viscosity on convective mixing behaviours could be captured simply by changing Ra, or if viscosity had an effect beyond a change in Ra.

3.2.3 Normalised interfacial area

The interfacial area, A, was noted to increase upon the onset of convection, shown by the triangle symbols on Figure 14. This was anticipated as the formation of convective fingers increases A.

The height of the first plateau was noted to increase with an increase in viscosity even though the systems in all three cases were within the same regime. Upon inspection of the 2D concentration distribution surface plots at the first plateau (Figure 15), it was noticed that while $Model_{inc}$ had fewer fingers than $Model_{dec}$, it had a more than proportional penetration depth which gave it a larger A overall before the system moved into the constant flux regime where merging occurred. This behaviour was due to the weaker convection currents in $Model_{inc}$, which allowed a greater penetration depth before merging occurred.

Furthermore, the maximum A was observed to vary inversely with viscosity. This could be explained by the varying convection strengths among the models. With a stronger convective term (similar to a higher Ra), $Model_{dec}$ could form smaller and sharper fingers, thereby giving rise to the highest A amongst the cases.

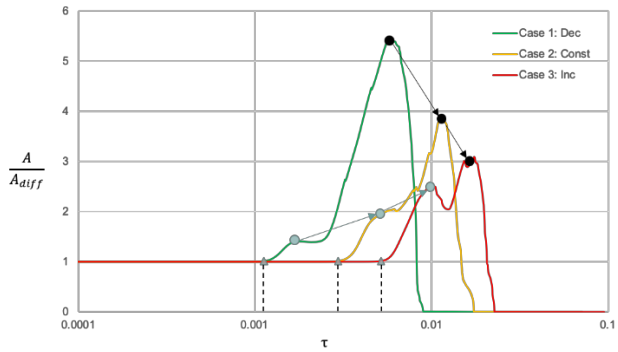


Figure 14. *Normalised interfacial area against dimensionless time*

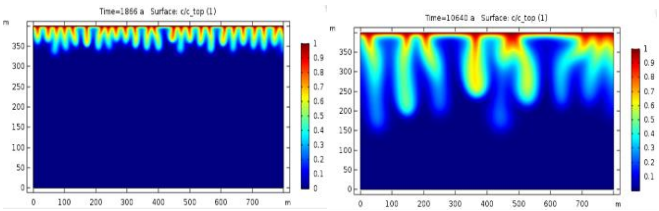


Figure 15. *Concentration distributions at first plateau for: left: $Model_{dec}$, right: $Model_{inc}$.*

4. Extended Discussion

An extended study, with motivations described in Sections 3.2.1 and 3.2.2, was performed to determine whether temporal changes in viscosity gave rise to effects which could be captured simply by a change in Ra or whether effects beyond a change in Ra would emerge.

The Ra value used to describe $Model_{dec}$ to $Model_{inc}$ thus far has been based on μ_0. However, μ changes over time for $Model_{dec}$ and $Model_{inc}$. Hence, a time- and space-weighted average viscosity, $\overline{|\mu|}$, was calculated to obtain a better estimate of the viscosity for each system. This value was then used to calculate an updated Ra, via the relation below (refer to Appendix G for calculations).

$$Ra = \frac{\kappa \Delta \rho g H}{\mu_0 \phi D} \rightarrow Ra_{updated} = \frac{\kappa \Delta \rho g H}{\overline{|\mu|} \phi D}$$

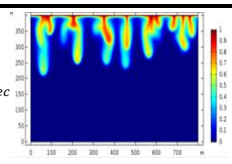
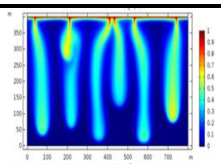
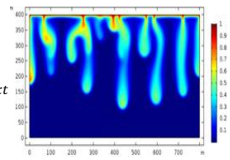
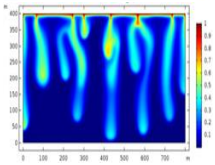
The evaluated $\overline{|\mu|}$ and $Ra_{updated}$ values for both models are shown in Table 5.

Table 5. *$\overline{|\mu|}$ and $Ra_{updated}$ values for $Model_{dec}$ and $Model_{inc}$*

Model	$\overline{\lvert\mu\rvert}$ / x 10^{-4} (Pa s)	$Ra_{updated}$
$Model_{dec}$	8.87	2000
$Model_{inc}$	24.0	740

For this study, emphasis was placed on $Model_{dec}$ as it was of greater interest for applications in depleted oil reservoirs. Hence, a comparison between $Model_{dec}$ and a new constant viscosity model with a Ra value of 2000 ($Model_{ext}$) was conducted.

Table 6. *Comparison of concentration distribution of $Model_{dec}$ and $Model_{ext}$ at constant τ and constant penetration depth*

Case	$\frac{c}{c_{top}}$ at $\tau = 3.24\ x\ 10^{-3}$	$\frac{c}{c_{top}}$ at $400m$	τ/ x 10^{-3} at $400m$
$Model_{dec}$			5.57
$Model_{ext}$			3.87

The 2D concentration distribution plots comparing the cases at constant τ and penetration depth are summarised in Table 6. As evidenced from the table, notable differences exist between the two cases.

Comparing across the same τ, $Model_{dec}$ was observed to show higher $scCO_2$ concentrations throughout the finger than $Model_{ext}$ while fading of the fingers midway was observed for $Model_{ext}$. These observations were likely because the fingers in $Model_{dec}$ have a lower viscosity than the bulk fluid. This posed less resistance within the fingers and facilitated the downward movement of the denser $scCO_2$-rich mixture from the top boundary to the fingertips. Another key observation was that $Model_{ext}$ had a deeper penetration depth than $Model_{ext}$. This could be accounted for as it possessed an initial edge; at $\tau = 0$, $Model_{ext}$ had a lower viscosity and thus higher Ra, which implied that it had a stronger convective term for finger growth.

An examination of the results for a constant penetration depth of 400m showed that $Model_{dec}$ had more defined vertical channelling of the less viscous fluid, which bypassed larger areas of the more viscous domain. This could be attributed to the lower viscosity of the mixture, which allowed for increased mobility. This, coupled with a downward force brought about by convection, aided the downward motion of the less viscous fluid. More fingertip splitting was also observed for $Model_{ext}$, which had a lower viscosity of $8.87 \times 10^{-4}\ Pa\ s$ compared to $1.44 \times 10^{-3}\ Pa\ s$ in $Model_{dec}$ at a depth of 400m. A plausible reason for the observed fingertip splitting is that a lower viscosity facilitates greater fluid movement, disrupting local equilibriums and inducing instabilities at the finger boundaries, which gave rise to fingertip splitting. Finger shielding was also noticed in $Model_{ext}$ where the growth of a more dominant finger stunted the growth of another.

Overall, the extended study demonstrated that changing the value of Ra was insufficient to capture the full effect of viscosity on convective mixing. The main differences lay in the finger patterns observed.

5. Conclusions and Outlook

In this study, the effect of temporal changes in mixture viscosity on the convective mixing of $scCO_2$ in oil reservoirs was numerically investigated through the successful development and implementation of a 2D Rayleigh-Bénard convection model in COMSOL®.

The results showed that for an initial Ra of 1000, the rate of mixing had an inverse relationship with viscosity. This was evident from: (i) the increasing finger penetration depths in the 2D surface plots at a specific time, (ii) the earlier onset time for convection, and (iii) the increasing maximum dimensionless flux observed when viscosity was decreased. It was also noted that quicker mixing did not compromise mixing performance. In addition, increasingly defined channelling, which resulted in higher interfacial areas, was noticed with a decrease in viscosity. The above results were noted to be similar to the effects of varying Ra without a change in viscosity.

An extended study conducted to establish whether the effect of a temporal change in viscosity contrast could be captured by simply increasing the Ra revealed unreconcilable differences in finger patterns. Therefore, a temporal change in viscosity contrast gave non-negligible effects – beyond changing Ra, it caused differences in the formation of finger patterns. These should be accounted for in future works involving the convective mixing of $scCO_2$ in oil, and especially for applications in geological carbon sequestration in depleted oil reservoirs.

This work serves as a basis for future work which could explore a Ra value of 10000 given that depleted oil reservoirs were found to have Ra ranging between 1000 to 10000 (see Section 2.2). A linear viscosity-concentration correlation could also be investigated in place of the exponential correlation used in this study since fitting a linear trendline to literature data (refer to Figure 2) gave an R^2 value of 0.789, which was within a 5% difference of that for the decreasing exponential curve. For added realism, a 3D model could be studied.

While the numerical results were credible, experimental validation would be desirable. This would involve utilising X-ray Computed Tomography scanning or analogue fluid systems experiments such as those done by Liyanage *et al.* (2018) and Neufeld *et al.* (2010).

For broader applications, this work could be adapted to study the simultaneous injection and dissolution of $scCO_2$ in oil (Tran, n.d.) during Enhanced Oil Recovery (EOR), which is allows for additional oil recovery (Offshore Technology, 2018). A background flow could be included into the current model to factor in the horizontal velocities associated with CO_2 injection.

Acknowledgements

We would like to thank Dr. Ronny Pini and Anna-Maria Eckel for their invaluable support, guidance and insight throughout the course of the project.

Supplementary Information

A separate document containing the raw literature data, sample calculations, raw numerical results from COMSOL®, 2D concentration distribution surface plots from COMSOL® and access links to the COMSOL® and MATLAB® files is enclosed.

References

BMT Fluid Mechanics (2010), "Computational Fluid Dynamics", available at: https://www.bmtfm.com/media/4127750/computational_fluid_dynamics.pdf (accessed 28 May 2019).

Calabrese, C. (2017) *Viscosity and Density of Reservoir Fluids with Dissolved CO_2*. Imperial College London.

CCSA (2016) 'What is CCS? – The Carbon Capture & Storage Association (CCSA)', *What is CCS? – The Carbon Capture & Storage Association (CCSA)*. Available at: http://www.ccsassociation.org/what-is-ccs/ (Accessed: 11 December 2019).

Celik, I.B., Ghia, U., Roache, P.J. and Freitas, C.J. (2008), "Procedure for Estimation and Reporting of Uncertainty Due to Discretization in CFD Applications", *Journal of Fluids Engineering*, Vol. 130 No. 7, p. 78001.

Ciotta, F. *et al.* (2009) 'Viscosity and Density of Carbon Dioxide + 2,6,10,15,19,23-Hexamethyltetracosane (Squalane)', *Journal of Chemical Engineering*, 54, pp. 2436–2443.

COMSOL (no date) 'Introduction to COMSOL Multiphysics'.

COMSOL Blog (2013), "Solutions to Linear Systems of Equations: Direct and Iterative Solvers", available at: https://www.comsol.de/blogs/solutions-linear-systems-equations-direct-iterativesolvers/ (accessed 30 April 2019).

De Wit, A. and Homsy, G. M. (1997) 'Viscous fingering in periodically heterogeneous porous media. I. Formulation and linear instability', *Journal of Chemical Physics*, 107(22), pp. 9619–9628. doi: 10.1063/1.475258.

Eckel, A.-M. (2019) *Three-dimensional convective mixing in geologic carbon sequestration applications*. Imperial College London.

Emami-Meybodi, H. *et al.* (2015) 'Convective dissolution of CO_2 in saline aquifers: Progress in modeling and experiments', *International Journal of Greenhouse Gas Control*. Elsevier Ltd, 40, pp. 238–266. doi: 10.1016/j.ijggc.2015.04.003.

Gallaire, F. and Brun, P. T. (2017) 'Fluid dynamic instabilities: theory & application to pattern forming in complex media', *Philosophical Transactions of the Royal Society A*. The Royal Society Publishing, 375(2093). doi: 10.1098/rsta.2016.0155.

Garcia, J.E. (2001), "Density of aqueous solutions of CO2", Graduate Student Research Assistant, Earth Sciences Division, Lawrence Berkeley National Laboratory, 2001.

Gershenzon, N.I., Ritzi, R.W., Dominic, D.F., Soltanian, M., Mehnert, E. and Okwen, R.T. (2015), "Influence of small-scale fluvial architecture on CO 2 trapping processes in deep brine reservoirs", *Water Resources Research*, Vol. 51 No. 10, pp. 8240–8256.

Ghesmat, K., Hassanzadeh, H. and Abedi, J. (2011), "The effect of anisotropic dispersion on the convective mixing in long-term CO2 storage in saline aquifers", *AIChE Journal*, Vol. 57 No. 3, pp. 561–570.

Hayhoe, K. *et al.* (2017) 'Ch. 4: Climate Models, Scenarios, and Projections', *Climate Science Special Report: Fourth National Climate Assessment*. Washington, DC, I, pp. 133–160. doi: 10.7930/J0WH2N54.

Hidalgo, J.J. and Carrera, J. (2009), "Effect of dispersion on the onset of convection during CO2 sequestration", *Journal of Fluid Mechanics*, Vol. 640, p. 441.

Intergovernmental Panel on Climate Change (IPCC) (2014) *Climate Change 2014 Synthesis Report Summary for Policymakers*, *IPCC 5th Assessment Report (AR5)*.

Javaheri, M., Abedi, J. and Hassanzadeh, H. (2009), "Onset of Convection in CO Sequestration in Deep Inclined Saline Aquifers", *Journal of Canadian Petroleum Technology*, Vol. 48 No. 08, pp. 22–27.

Kitanidis, P. K. (1994) 'The concept of the dilution index', *Water Resources Research*, 30(7), pp. 2011–2026.

Kwaśniewski, L. (2013), "Application of grid convergence index in FE computation", *Bulletin of the Polish Academy of Sciences: Technical Sciences*, Vol. 61 No. 1, pp. 123–128.

Land Trust Alliance (2018) *Carbon Dioxide, Methane, Nitrous Oxide, and the Greenhouse Effect - Conservation in a Changing Climate*. Available at: https://climatechange.lta.org/get-started/learn/co2-methane-greenhouse-effect/ (Accessed: 11 December 2019).

Lindeberg, E. and Bergmo, P. (2002), "The Long-Term Fate of CO2 Injected into an Aquifer", *Greenhouse Gas Control Technologies - 6th International Conference*, I, pp. 489–494.

Liyanage, R. *et al.* (2018) 'Multidimensional Observations of Dissolution-Driven Convection in Simple Porous Media Using X-ray CT Scanning', *Transport in Porous Media*. Springer Netherlands, 126(2), pp. 355–378. doi: 10.1007/s11242-018-1158-3.

Nazari Moghaddam, R., Rostami, B. and Pourafshary, P. (2015), "Scaling Analysis of the Convective Mixing in Porous Media for Geological Storage of CO 2 An Experimental Approach", *Chemical Engineering Communications*, Vol. 202 No. 6, pp. 815–822.

Nazari Moghaddam, R., Rostami, B., Pourafshary, P. and Fallahzadeh, Y. (2012), "Quantification of Density-Driven Natural Convection for Dissolution Mechanism in CO2 Sequestration", *Transport in Porous Media*, Vol. 92 No. 2, pp. 439–456.

Neufeld, J.A., Hesse, M.A., Riaz, A., Hallworth, M.A., Tchelepi, H.A. and Huppert, H.E. (2010), "Convective dissolution of carbon dioxide in saline aquifers", *Geophysical Research Letters*, Vol. 37 No. 22, n/a-n/a.

Newell, D.L., Carey, J.W., Backhaus, S.N. and Lichtner, P. (2018), "Experimental study of gravitational mixing of supercritical CO 2", *International Journal of Greenhouse Gas Control*, Vol. 71, pp. 62–73.

OECD (2012) 'OECD ENVIRONMENTAL OUTLOOK TO 2050: What could the environment look like in 2050?'

Offshore Technology (2018) *Why is enhanced oil recovery becoming an essential technique?* Available at: https://www.offshore-technology.com/comment/enhanced-oil-recovery-becoming-essential-technique/ (Accessed: 9 December 2019).

Pau, G.S.H., Bell, J.B., Pruess, K., Almgren, A.S., Lijewski, M.J. and Zhang, K. (2010), "Highresolution simulation and characterization of density-driven flow in CO2 storage in saline aquifers", *Advances in Water Resources*, Vol. 33 No. 4, pp. 443–455.

Perelmuter, A. v. and Slivker, V.I. (2003), *Numerical Structural Analysis: Methods, Models and Pitfalls, Foundations of Engineering Mechanics*, Springer Berlin Heidelberg, Berlin, Heidelberg, s.l.

Rauwendaal, C. and Tucker, C.L. (1991), Mixing in polymer processing: *Principles of Mixing Measurement*, M Dekker, New York.

Riaz, A. and Cinar, Y. (2014) 'Carbon dioxide sequestration in saline formations: Part I-Review of the modeling of solubility trapping', *Journal of Petroleum Science and Engineering*. Elsevier, 124, pp. 367–380. doi: 10.1016/j.petrol.2014.07.024.

Simmons, C.T., Fenstemaker, T.R. and Sharp, J.M. (2001), "Variable-density groundwater flow and solute transport in heterogeneous porous media: approaches, resolutions and future challenges", *Journal of Contaminant Hydrology*, Vol. 52 No. 1-4, pp. 245–275.

Smith, P. *et al.* (2016) 'Biophysical and economic limits to negative CO2 emissions', *Nature Climate Change*, pp. 42–50. doi: 10.1038/nclimate2870.

Tang, Y. *et al.* (2019) 'Experimental Study on the Density-Driven Carbon Dioxide Convective Diffusion in Formation Water at Reservoir Conditions', *ACS Omega*. American Chemical Society, 4(6), pp. 11082–11092. doi: 10.1021/acsomega.9b00627.

Tran, J. (no date) *Enhanced Oil Recovery | Student Energy*. Available at: https://www.studentenergy.org/topics/enhanced-oil-recovery (Accessed: 9 December 2019).

United Nations Climate Change (UNCC) (2015) *The Paris Agreement*. Available at: https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement (Accessed: 11 December 2019).

Wang, L., Nakanishi, Y., Hyodo, A. and Suekane, T. (2017), "Three-dimensional Finger Structure of Natural Convection in Homogeneous and Heterogeneous Porous Medium", *Energy Procedia*, Vol. 114, pp. 5048–5057.

Zhang, Z. and Agarwal, R.K. (2014), "Numerical Simulation of CO_2 Sequestration in Large Saline Aquifers", in Esteves, V. (Ed.), *CO_2 Sequestration and Valorization*, InTech.

Zimmerman, W. B. and Homsy, G. M. (1992) 'Three-dimensional viscous fingering: A numerical study', *Physics of Fluids A*, 4(9), pp. 1901–1914. doi: 10.1063/1.858361.

Alcohol-Sensitive Holographic Sensor

Yimeng Jiao and Menghan Liu
Department of Chemical Engineering, Imperial College London, U.K.

Abstract An alcohol-sensitive holographic sensor was fabricated by recording reflection hologram gratings into a poly-2-hydroxyethyl methacrylate film using a Q-smart 850 pulsed Nd:YAG laser with wavelength of 355 nm, at energy setting 100 mJ. Changes in holographic diffraction peak wavelength in response to different alcohol analytes is characterised with a spectrophotometer for concentration ranges from 0 to 100 vol% in deionised water solution. Evaluation of spectral results show that the sensor can be used for short chain alcohols concentrations up to 60 vol%. The reversibility of the sensor is demonstrated and exhibits an average stabilisation time of 20 minutes. Simulation of the fabrication process is conducted in MATLAB using 2-D finite-difference time-domain algorithm to model the interference pattern and energy flux profile of laser beam recording in the polymer medium. This work demonstrates the proof-of-concept of a nanoparticle-free holographic sensor giving continuous, reversible, and fast readouts that allows the quantification of a known alcohol as well as the identification of relative alcohol concentrations in unknown mixtures.

Keywords: Holographic sensor; alcohol sensor; holographic grating simulation; finite-difference time-domain;

Introduction

In the beverage industry, prominent methods for quantifying alcohol concentration often include the use of expensive and highly analytical equipment such as modular distillation units (Buchi, 2019), and density-based systems (Anton Paar, 2019), which require frequent calibration and multi-sample collection, pre-treatment steps.

Meanwhile, optical sensors have become attractive for usage in biochemical systems, due to their ability to provide label-free quantification of designated external stimuli in real-time. Holography allows for the fabrication of optical devices, with the advantageous qualities including lightweight, low-cost, and disposable.

Previous works on alcohol concentration detection using holographic sensors involve a silver-halide based system, utilising silver nanoparticles (NP), similar to processes in photography (Mayes *et al.*, 1999). Thus, the possibility of manufacturing certain polymer matrix structures to monitor alcohol concentration through volumetric changes that alter the internal refractive index profile has been demonstrated.

In this work, a silver halide and NP-free sensor with holographic grating designed to provide colorimetric, real-time alcohol quantification was fabricated and characterised. In the synthesis step, a polymer backbone was constructed using P1 monomer solution, a mixture of HEMA with crosslinkers, Then, a second monomer solution (P2) with a higher crosslinker concentration was polymerised within the base polymer. This is known as a double polymerisation process.

The degree of crosslinking dictated the extent of volumetric change upon exposure to external stimuli, the analyte. The polymerisation reaction followed the periodic pattern of constructive and destructive interference pattern of the laser beam standing wave, thus creating diffraction gratings, which acted as a wavelength filter under the incident white light, and replayed a narrow band spectrum (Yetisen *et al.*, 2014). Therefore, the characterisation of the sensors was determined by the holographic diffraction wavelength, which varied upon swelling and shrinkage of the polymeric structure.

A simulation of the fabrication process was also conducted using MATLAB 2019b to further develop understanding in the laser writing process in polymer systems. It is known that the diffraction grating produces a spectral peak, i.e. the replay peak detected with optical fibre and sent to a colour processing software to produce a spectrum, at a wavelength which can be approximated by Bragg's law ($\lambda_{peak,reflection} = 2n\Lambda cos(\theta)$, where n is the average refractive index, Λ is the periodic fringe spacing, and θ is the angle of reading white light to the normal). This report aims to act as a basis to further works to investigate the resultant refractive index profile from the volumetric modulations in response to analyte carbon chain length and concentration.

The potential of the product to be further developed into a commercial sensor requiring low operational expertise, no recalibration, was demonstrated for methanol, ethanol, and propanol isomers; a reagent alcohol of known composition and butan-1-ol was also tested.

Methods

Equipment

Hologram construction: Microscope slides, a Single-side polyester coated aluminium sheet, dichroic mirrors, an uncoated plano-concave lens, Philips HB 172 solarium lamp, frequency tripled Q-smart 850 pulsed Nd:YAG laser with wavelength of 355 nm, energy of 100 mJ was used for the holographic recording, with a delay time of 10 µs, so the output energy was maximised.

Optical characterisation: A glass cutter, a glass cuvette, an Ocean Optics flame spectrometer connected to a computer with Oceanview software installed which converts colours into intensity vs wavelength plots in real-time and has a detection wavelength range between 350nm to 1000nm, optical fibres, a Tungsten halogen light source, a articulating base stage, a collimating lens and a smartphone for recording colorimetric responses.

Material: 3-(trimethoxysilyl) propyl methacrylate, acetone, 2-hydroxyethyl methacrylate (HEMA), glycerol dimethacrylate (GDM), DMSO, 2-hydroxy-2-methylpropiophenone(HMPP),ethylene dimethacrylate (EDMA), methanol, DI water, ethanol, propan-1-ol, isopropanol, butan-1-ol.

Pre-silanisation of slides: Clean glass slides were immersed in silane solution of 3-(trimethoxysilyl) propylmethacrylate and acetone with a volume ratio of 1:50 for 24 hours at room temperature, then rinsed with ethanol and excess solution was gently wiped off, where the slides were functionalised to allow better adherence of the polymer layer on the slide through covalent bonds formation of a siloxane network (-Si-O-Si-) (Yuk *et al*., 2015).

Monomer solution preparation: P1 solution containing HEMA 4 mL as the backbone material for the polymer, 0.3 mL of GDM as the crosslinker and 4 mL HMPP used for UV photoinitialiser. The monomer solution for P2 differed from P1 solution with different amount of HEMA (0.4 mL) and EDMA (4 mL) as the crosslinker, methanol 4 mL and DI water 0.8 mL.

Polymerisation: 2 strips of standard insulation tapes with the thickness of 0.178mm were adhered onto the Al side of a polyester coated aluminium sheet in parallel with a separation of around 70 mm (which was slightly less than the length of a glass slide), 350 µL of P1 monomer solution was pipetted in the middle and a silanised glass slide was slowly placed on the solution globule with the two ends lifted by the tapes from the sheet surface, thus, a relevantly uniform thickness of P1 polymer was obtained. Any bubbles should be avoided to prevent unwanted scattering, then exposed under UV light for 30 mins.

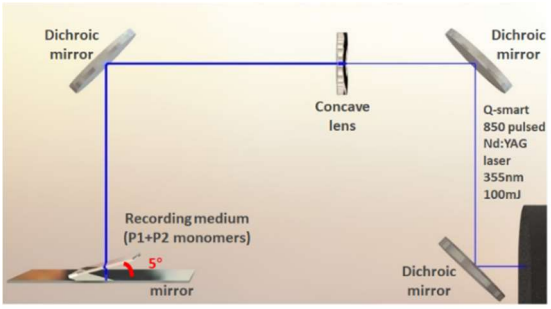


Figure 1. Laser exposure step experimental set up

After exposure, in darkened conditions, the slide was set on 350 µL of P2 monomer solution on the aluminium sheet for more than 5 mins, excess solution was then wiped off, and blow-dried with cool air. The laser was set at single pulse, 5 ns, 355 nm wavelength. The laser beam passed through the glass slide with the polymer side facing down, fell directly on the recording medium, whilst the redirected beam from a mirror beneath was scattered back onto the emulsion. Diffraction gratings were produced due to the constructive and destructive interference of two coherent beams.(see figure 1 for the construction setup) One side of the glass slide was elevated to an angle of 5° with respect to the mirrored surface, so that the diffraction wave was separated from the incident white light and its reflection with an angle in the later stage when the characterisation of the sensor was determined. The slide was then immersed in an ethanol solution (50 vol% in water) to remove unpolymerised monomers in the dark fringes. The hologram after exposure could be used to sense alcohol instantaneously.

Sensor characteristics measurements:
The tuneability of the sensor was analysed by detecting the spectral response of the sensor in reflection mode with an optical fibre connected to the spectrophotometer, under the exposure of a normal incident white light source, while the sensor(cut off from the glass slide to fit inside a cuvette with the polymer side facing the solution) was immersed in solutions of methanol, ethanol, propan-2-ol, propan-1-ol and butan-1-ol, as well as reagent alcohol (a mixture of 90% ethanol, 5% methanol, 5% propran-2-ol by volume), with 0 – 100 vol% concentrations in DI water, at room temperature, a constant wavelength was reached before any reading taken.
The time taken upon changes of the sensor was demonstrated through introducing alcohol solutions with concentrations of 20 vol%, then replaced with 70 vol% and back to 20 vol%.

Fabrication simulation

Modelling of laser beam

The diffraction grating was written onto the polymer medium via crosslinking polymerisation reactions, which were initialised by the standing wave generated from the counterpropagating laser beams, in the setup shown in figure 1. This process was simulated in MATLAB by modelling the laser source as travelling electromagnetic (EM) fields, where the primary incident beam combined with the reflected beam to form an interference pattern.
EM waves are with the Maxwell's curl equations: modified Ampere's law (*Eqn.(1)*) and Faraday's law (*Eqn.(2)*), both of which describe the inductive relationship between a time-varying field and a circulating field.

$$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} \qquad (1)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad (2)$$

Where $\vec{H}$ is the magnetic field, $\vec{J}$ the total current density, $\vec{D}$ the electric displacement field, $\vec{E}$ the electric field, $\vec{B}$ the magnetic flux density. To find a self-consistent solution for the EM field, Maxwell's equations require

the constitutive relations, as shown in *Eqn.(3)&(4)*, that describe the behaviour of materials under the influence of the fields, where ε is the dielectric constant (permittivity), in this work relative values (divided by free permittivity of free space) are used; similarly, μ is the relative magnetic permeability. These EM qualities of the polymer are assumed to remain constant throughout the laser writing process.

$$\vec{B}(t) = [\mu(t)] \cdot \vec{H}(t) \qquad (3)$$

$$\vec{D}(t) = [\varepsilon(t)] \cdot \vec{E}(t) \qquad (4)$$

MATLAB simulation

In this project, the dimension of the glass substrate on which the polymer medium was fabricated was 3 cm by 1 cm, where the thickness of the polymer is assumed to be 0.178 mm (the thickness of insulation tapes adhered on the surface of Al-coated polymer sheet). During spectral analysis, a segment of the polymer of dimension 2 cm by 0.5 cm was cut to use as holographic sensor due to the size constraint of the cuvette used to soak the sensor in analyte. From figure 2, one can see that the diffraction grating was generated along the thickness of the polymer medium.

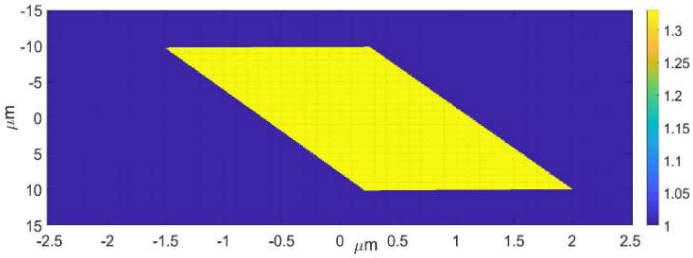


Figure 2. GUI set up. The colour map bar indicates the relative ε, i.e. blue for surroundings since it is air with value 1, and yellow for the polymer medium which is simplified as pure pHEMA with ε=1.77.This is a vertical rendition of the experimental setup; the polygon is tilted at an angle of 5° anticlockwise to the normal. The y-axis is the height and x-axis the thickness of the polymer medium. Two laser sources are defined on either side of the polygon with height 10 μm and width 0.1 μm. The four borders of the GUI window are set as PML to eliminate reflection effects; thus, the only interference occurs between laser source travelling in opposite directions along the x-axis.

Therefore, in MATLAB simulation, a 2-D model was constructed where the y-axis represented the height of the polymer, x-axis the thickness, and the polymer was assumed to exhibit uniform behaviour in the z direction, as shown in figure 2.

The simulation process is summarised by figure 3(a). The basis of EM wave modelling is the finite-difference time-domain (FDTD) method, which solves the coupled Maxwell's equations numerically (Yee, 1966). The DTD method starts by replacing the derivative terms in *Eqn.(1)&(2)* with finite differences. From this discretisation of space and time, and combination of the curl equations with their constitutive relations, "update equations" in time *Eqn.(5)&(6)* are obtained, with assumption that all materials are linear, non-dispersive, and isotropic, with no magnetic losses. Thus, the current term $\vec{J}$ is neglected from *Eqn.(1)*, and the EM parameters ε and μ are constants. These describe unknown 'future' field values in terms of known, initialised 'past' field values. It is important to note that for the Yee grid is defined such that for the same spatial indices, the H and E fields are staggered as shown in figure 3(b), therefore the H field is half a spatial step further than the E field at the same indices, resulting in $\vec{E}(t)$ corresponding to $\vec{H}\left(t + \frac{\Delta t}{2}\right)$.

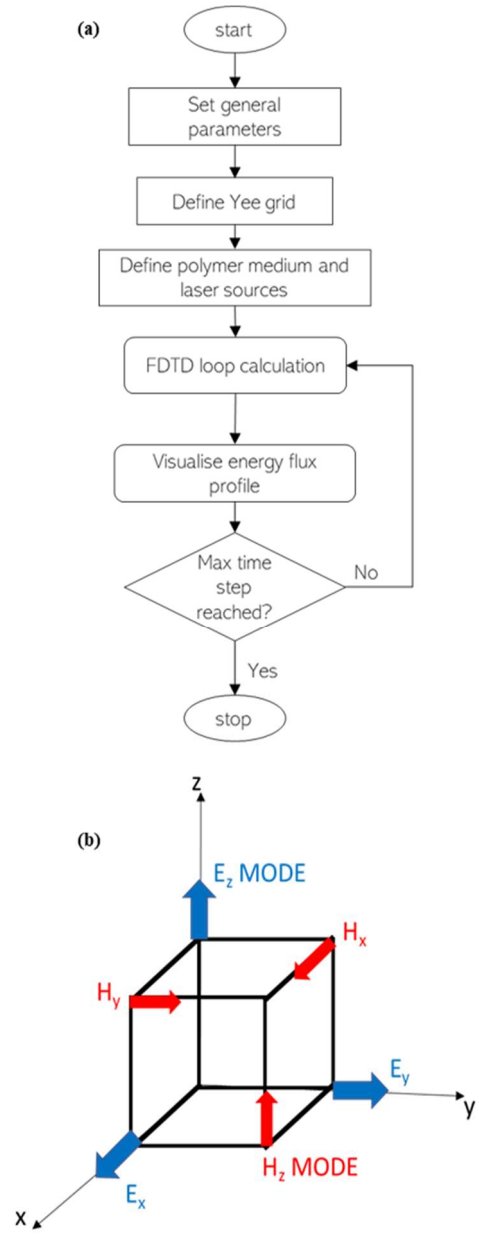


Figure 3. (a) Flow chart of numerical simulation of laser exposure in fabrication process. (b) To compute the FDTD algorithm in 2-D, we start from the 3-D model and simplify by eliminating the derivatives in the z-direction. To illustrate the staggering of E and H fields, a 3-D Yee cell is presented here. In 2-D FDTD update equations, there are two modes: the E_z mode, which updates H_y and H_x in terms of E, and the H_z mode, which updates E_x and E_y in terms of H. These two modes update field values at the same spatial point, but as shown in this figure, the positions of the field components seem to be shifted, i.e. staggered.

Meanwhile, the spatial derivatives only reside in the curl equations, and in 2-D their update equations in space are depicted by *Eqn.(7)&(8)*.

$$\vec{H}\left(t+\frac{\Delta t}{2}\right)=\vec{H}\left(t-\frac{\Delta t}{2}\right)-\frac{\Delta t}{\mu}\vec{\nabla}\times\vec{E}(t) \qquad (5)$$

$$\vec{E}(t+\Delta t)=\vec{E}(t)+\frac{\Delta t}{\varepsilon}\vec{\nabla}\times\vec{H}\left(t+\frac{\Delta t}{2}\right) \qquad (6)$$

$$\frac{\vec{E}_y^{new}-\vec{E}_y^{old}}{\Delta x}-\frac{\vec{E}_x^{new}-\vec{E}_x^{old}}{\Delta y}=\mu\vec{H}_z \qquad (7)$$

$$\frac{\vec{H}_y^{new}-\vec{H}_y^{old}}{\Delta x}-\frac{\vec{H}_x^{new}-\vec{H}_x^{old}}{\Delta y}=\varepsilon\vec{E}_z \qquad (8)$$

Prior to simulation, initial parameters are defined. Boundary conditions at the simulation border are chosen to be perfectly matching layers (PML), which absorb all radiation, as it is assumed that reflection from surroundings is negligible. Permittivity and permeability of background environment is chosen as that of air. Size dimension of x and y axis are chosen to be 5 μm and 30 μm respectively, to accommodate the scaled down model of the polymer segment used as sensor (see supplementary material for detail).

After the definition of the general parameters for this simulation, the graphical user interface (GUI) appears, and the polymer is constructed as a polygon with EM properties as that of pure pHEMA, since this is the bulk material. In addition to the material and PML information, the FDTD algorithm requires: the source wave length (355nm); angle of incidence, which is handled manually by the construction of the polymer polygon; the mode of polarisation, chosen as E_z, which serves as the initialisation of source electric field value ; and finally, the grid resolution, which is maximised along with the number of time steps, subject to computing memory constraints. Once all parameters are fully defined, they are overlaid onto a grid made of unit cells as shown in figure 2. Then matrices are constructed for the curl derivative values, into which E and H field values looped over all x and y direction are calculated and stored. This is done for each time step. The FDTD cycles run until the maximum number of time steps is reached; the optimum time step is found through an iterative process to ensure the simulation only terminates when the standing wave is fully developed, and at a sufficiently high resolution.

Results and Discussion

Material selection

Based on the paper published in 1999 by A. Mayes, HEMA was chosen as the polymer backbone material because of its more hydrophobic property, a relevantly large expansion of the polymer matrix in alcohol/water mixture would be observed and covering a wider range of alcohol concentrations; comparing to other more hydrophilic materials with the same ratio of crosslinker added （EDMA）, such as polyacrylamide, the resulting hologram not only would adhere poorly on the substrate but also merely respond to a small range of concentrations with a small extent of contraction.

The solvent DMSO in P1 monomer solution acted as the porogen, which generated pores within P1 matrix and allowed P2 monomers to enter the vacant space and DMSO was washed off after the polymerisation of P1 monomers.

HMPP was the photoinitiator; its role was to trigger the free radical reaction. Previous works used DMPA for photoinitiated crosslinking polymerisations; however, HMPP is shown to have a more hydrophilic nature, with a higher absorbance at the construction laser wavelength (Sigma Aldrich, 2019). Therefore, HMPP was used, as it would have higher mobility in P2 polymerisation.

The amount of crosslinker present in the polymer plays a significant role in the refractive index variation upon swelling for the sensor, a larger difference in refractive index between the where P2 was polymerised at and the baseline polymer P2 would result in a brighter hologram. And the maximum degree of swelling reduces with more crosslinker concentration in the polymer. Highest possible crosslinker concentration was needed for the P2 and as little crosslinker as possible for P1 while the integrity of the grating structure was maintained by crosslinking between polymer chains, thus the ratio was carefully selected through practical work.

In addition, due to the difference in hydrophilicity of the monomer solutions, different crosslinking agents were employed in P1 and P2. Since P1 is more hydrophobic overall, the use of a more hydrophilic crosslinker would allow more P2 monomers to be taken up by the P1 polymer structure. In the P2 monomer solution, EDMA was used; it is less hydrophilic than the GDM used in P1 monomer solution. GDM has an additional hydroxy group, thus EDMA would interact more with the hydrophobic P1 polymer matrix. The more P2 monomers enter P1, greater area of more highly crosslinked polymerisation would take place in the bright fringes of interference pattern. Methanol also aids P2 molecules entering P1 monomer matrix, due to its high water mobility; it was evaporated easily due to its low boiling point before laser exposure.

Alcohol-sensitive Hologram fabrication

The fabrication process detailed in this work consists of coating a layer (~0.178mm thick) of primary polymer solution (P1) onto a silanised glass slide, before immersing the slide into a secondary polymer solution (P2). The double-polymerised slide was then exposed to a laser beam and the redirected beam from a mirror placed beneath the slide, producing a polymer film containing an interference fringe pattern comprised of strongly and weakly crosslinked chains with a spacing of $\lambda/2$ along the thickness of the polymer medium. This is immediately usable as an optical sensor, whereby

illumination with a white light source recreates a monochromatic visual replay of the light wave by diffraction. The resulting spectral peak from diffraction gratings of hologram follows Bragg's law.

When the sensor is immersed in analyte solution, molecules enter the polymer matrix and cause swelling. Each periodic region has different capacity to swell due to the diffraction grating, and the extent of swelling is directly affected by the number of target analyte molecules. In this way, a reflection hologram is generated where the replay colour is dependent on the concentration of the analyte sample.

To ensure the replay wavelength remaining within the visible spectrum to analyse the colorimetric response of the sensor, the blow-drying step in the evaporation of methanol and DI water step during fabrication process was conducted on a low heating setting. This caused the polymer to contract so that without any degree of swelling, the diffracted narrow band spectrum would centre at its construction wavelength (Benton, 2008). This ensures that the shortest replay wavelength while swelling upon immersion in the tested analytes, which occurs when the sensor is exposed to pure water, is at 385nm, (which could be seen in figure 4(b), the spectra obtained for butan-1-ol) since water-alcohol mixture showed higher ability of swelling up the polymer than DI water alone and would only replay longer diffracted peak wavelengths (Mayes *et al*., 1999).

An alternative constructive wavelength of 532 nm was used for the construction of pHEMA based alcohol sensor by Mayes, however, the replay wavelengths may exceed the detection range for the spectrophotometer when the polymer is highly swelled. Another reason against using 532 nm laser was that the concentration range that gives a colorimetric response is much limited, since the visible light wavelength is from about 380 to 740 nm, and for higher concentrations, the replay wavelength would remain in the UV region.

Incident light spectrum

The spectrum of the white reading light source is shown in figure 4(a). From the waveform, it is clear that signal intensity is approximately 5 times greater in the green visible light region than in near UV. This would cause a proportional difference in diffracted light intensity at the peak wavelength of each analyte concentration. Since only the colorimetric response is used in analyte quantification, the spectral response was normalised as the relative intensity between the peaks are insignificant.

As seen in figure 5, the wavelength shift, read from plots (e)-(g) y-axis, refers to the change in peak wavelength relative to its peak position in comparison with the initial response of the sensor upon the immersion in pure DI water. The figure also shows that for all alcohol species, there is a linear relationship between $\Delta\lambda_{peak}$ and analyte concentration up until around 40 vol%, which is an improvement from the 30 vol% range of previous study on the polyacrylamide based alcohol holographic sensor (Mayes *et al*., 1999).

In addition, the colorimetric response (figure 4(a)-(d)) are very distinct, thus, the holographic sensor can provide a good, easily distinguishable visual readout of alcohol concentration in real time.

Butan-1-ol

Due to the low solubility of butan-1-ol in water, it was not immiscible above 10 vol%, therefore low concentrations were tested from 0 vol% to 8 vol%, with their spectra shown in figure 4(b) with intensity following the shape of the incident light spectrum, in addition, from figure 4(a), the linear relationship between peak wavelength shift and analyte concentration is particularly evident for the small butan-1-ol increments. This further shows that the sensor response is linear for concentration ranges 0 to 40 vol%, which allows the holographic replay to be used as an accurate alcohol quantifier.

Since other alcohols are completely miscible with water, larger increments in concentration were tested, from 0 - 100 vol%.

Propan-1-ol

Figure 4(c)-(d) are the results after noise removal for propan-1-ol. As expected, peaks were obtained at longer wavelength for higher concentrations up to 60 vol%. However, the intensity appeared to not increase with concentration, which is not consistent with the intensity profile of incident white light spectrum from figure 4(a). This is likely due to a pre-set low integration time in the Oceanview processing software. It is related to the level of background noise. To observe peaks at lower wavelengths, the resolution of peaks at higher concentrations must be compromised. Therefore, only the positions of the peaks are significant in figure 5(a)-(d).

Interestingly, the peak wavelength decreases above 60 vol%, as shown by the curves moving towards shorter wavelength with increasing propan-1-ol concentration in figure 4(d). This behaviour is consistent across all alcohols and does not affect the reversibility of the sensor – as shown later in figure 5(e)-(g). Also, it is noticed that bell-shaped curves are obtained, whose peak wavelength shifts at concentrations of 70 vol% to 80 vol% do not vary as significantly as for lower concentrations. This may be due to the sensor becoming saturated, as the polymer matrix expels solution and contracts with increasing concentration. The cross-links between polymer chains limit the extent of expansion. However, the decrease of λ_{peak} for concentrations above 60 vol% could also be due to several factors. Firstly, the interaction between analyte solution and the polymer matrix, mainly through hydrogen bonding, might change. Previous experimental study observed that water-alcohol mixtures show dramatic variations in hydrogen bonding in the water-rich and alcohol-rich region, on thermodynamics and structures (Ballal and Chapman, 2013), leading to fewer molecules penetrate into the matrix at higher concentrations; secondly, it is

possible that the osmotic mechanism draws limited solution into the matrix to balance out the pressure difference with the bulk solution. Further study is still required to improve understanding in this aspect.

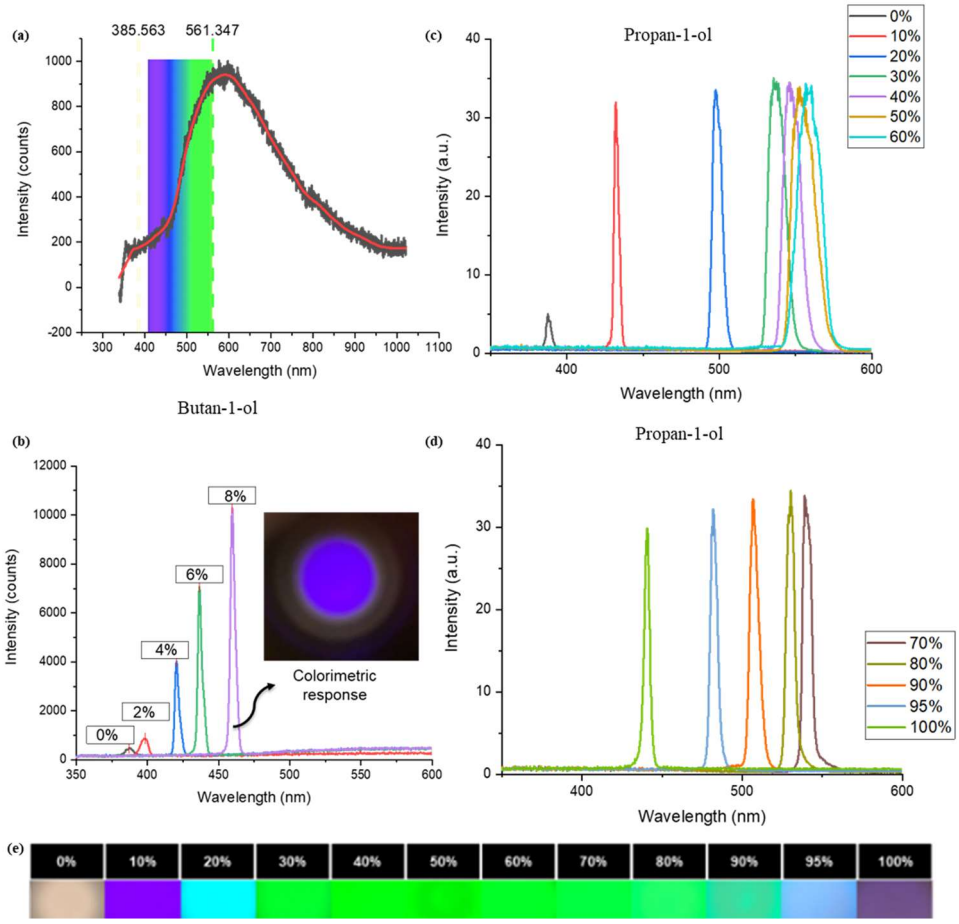


Figure 4. Sensor responses. (a) White light spectrum and the range of replay wavelength. The black waveform is the raw data, the red line is the smoothed signal. The visible colorimetric response is shown as the colour block from violet to green. (b) Butanol raw data obtained from the Oceanview software. (c) Propan-1-ol after noise removal for concentrations 0 vol% - 60 vol%. (d) Propan-1-ol after noise removal for concentrations 70 vol% - 100 vol%. (e) Time taken upon changes as a function of methanol (black data points), ethanol (red), propan-1-ol (blue), popan-2-ol (green), where the sensor is initially stable in analyte solution of 20 vol%, and replaced with 70 vol%, then back to 20 vol%.

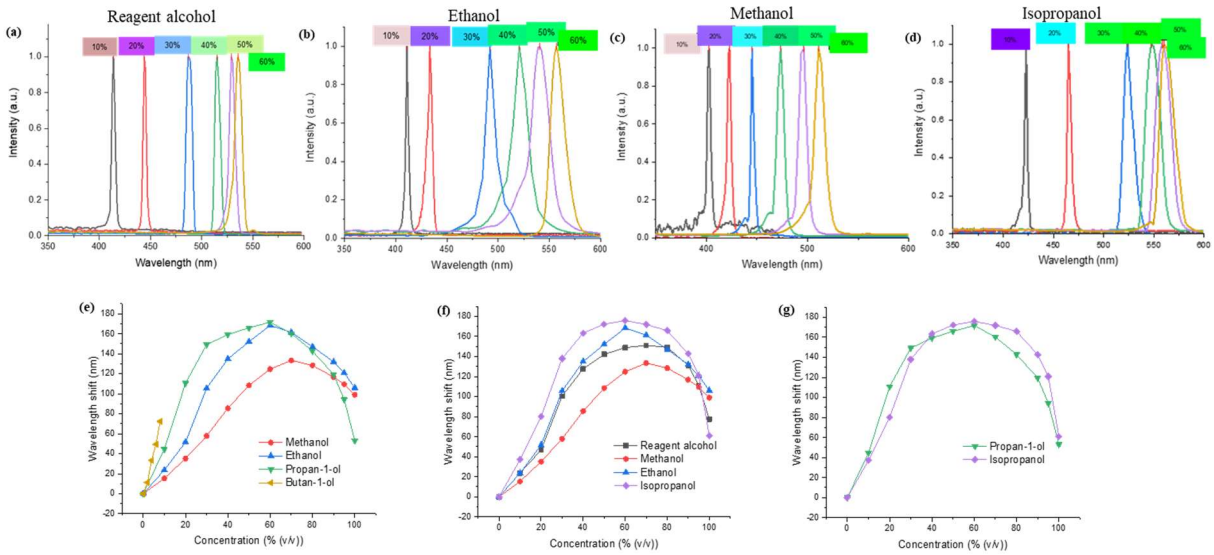


Figure 5. (a)-(d) Normalised and smoothed spectra as a function of wavelength for (a) reagent alcohol; (b) ethanol; (c) methanol; (d) Isopropanol for solutions with concentrations of 0 vol% to 60 vol%.
(e)-(g) Comparison between diffracted peak wavelength for different alcohol species.

Reagent alcohol and its component alcohols

Figure 5(a)-(d) illustrates the normalised spectra of all the other alcohols between concentration of 10 – 60 vol%, it can be seen that for reagent alcohol, ethanol and methanol at 10 vol% the sensor responses are near UV (mostly white reflection were recorded in the images); however, for isopropanol, a bright blue colour was observed, indicating a much greater extent of swelling compared to the other alcohols at the same concentration. Since it is known that reagent alcohol contains 90% ethanol, 5% methanol and 5% isopropanol by volume, by comparing the curves of reagent alcohol and ethanol in figure 5(f), it can be concluded that methanol has a more significant effect since reagent alcohol provides lower wavelength shift than ethanol overall. This shows that the sensor can be used to identify the relative concentrations in a mixture of different alcohols, which can be an interesting platform to explore for industrial processes where the production rates of multiple alcohols need to be monitored.

Propan-1-ol vs propan-2-ol

Propan-1-ol and propan-2-ol are isomers, which produce similar extents of swelling as seen in figure 5(g); at concentrations under 40 vol%, propan-1-ol exhibits slightly higher $\Delta\lambda_{peak}$. It is known that Propan-1-ol has a longer linear hydrocarbon chain, and is thus more hydrophobic than propan-2-ol. Consequently, it has higher hydrophobic interaction with the polymer matrix, resulting in slightly more significant wavelength shifts over the low concentrations. One can see that significant deviations only occur at higher concentrations, as shown by the larger gap between the curves after 40 vol%, where an opposite effect was observed, where $\Delta\lambda_{peak}$ is lower for the primary alcohol. It is speculated that this is due to the complex water-alcohol mixture property changes as aforementioned.

Effects of alkyl chain length of alcohols

The ability of each alcohol to cause swelling in the matrix at lower concentrations increases with alkyl chain length, as shown in figure 5(e). This could be attributed to the fact that, for same number of analytes interacting with the polymer matrix, due to the extended hydrocarbon chain length, there is a greater hydrophobic interaction which could generate a higher degree of swelling.

All the tuning experiments were carried out on the same sensor. The sensor demonstrated reversibility by returning to the original state after a period of stabilising as shown in figure 6, where the $\Delta\lambda_{peak}$ for each analyte concentration remains the same after the sensor is exposed to a different concentration.

The slope of the curves shows the response speed of the sensor upon analyte concentration changes; comparison between the unbranched primary alcohols, methanol, ethanol and propan-1-ol, show that the response speed decreases with increasing alkyl chain length of alcohols, which is in contrast with the trend for maximum degree of polymer swelling. The shift in wavelength for methanol solutions has the steepest slope, showing it is the fastest to stabilise. Propan-1-ol, however, takes the longest to reach equilibrium. This could be due to the effects of size and polarity. Methanol is the smallest molecule and has the highest polarity. Therefore, it enters the matrix and causing the swelling faster.

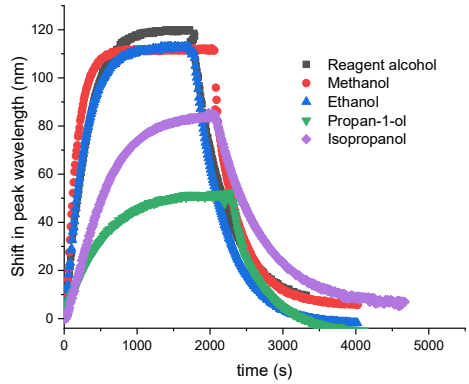


Figure 6. Sensor responses: (a) Time taken upon changes as a function of methanol (black data points), ethanol (red), propan-1-ol (blue), popan-2-ol (green), where the sensor is initially stable in analyte solution of 20 vol%, then replaced with 70 vol% and back to 20 vol%.

On the other hand, propan-2-ol exhibits a speed nearly double than that of propan-1-ol; this could be caused by the different shapes of analyte molecules in comparison with the pore size: propan-2-ol is a branched chain hydrocarbon (the hydroxy group is in the centre of the hydrocarbon chain). The secondary alcohol is bulkier and more sphere-like, compared to the straight chain propan-1-ol. As aforementioned, the pores in P1 matrix was generated by the space taken up by DMSO molecules which were later washed off. Since DMSO is a trigonal pyramidal molecule (Thomas *et al.*, 1966), it has a shape that is more similar to propan-2-ol, therefore shape selectivity might cause the slower uptake of more linear shaped molecules, which are then more likely to be hindered while entering the matrix. Interestingly, the relative polarity of propan-2-ol is slightly lower than propan-1-ol; this could suggest that the effect of shape selectivity outweighs the effect of polarity.

Figure 6 also shows that the average response time is 20 minutes for the four different types of alcohols. This was longer than expected, since previous pHEMA-based holographic sensor required an average of 50 seconds to reach equilibrium (Yetisen *et al.,* 2014). However, experimental parameters such as the target analyte, crosslinker, polymer medium thickness, were all different for this work. Therefore, one can infer that holographic sensors can have a large range of possible response time depending on fabrication conditions.

Simulation of fabrication process

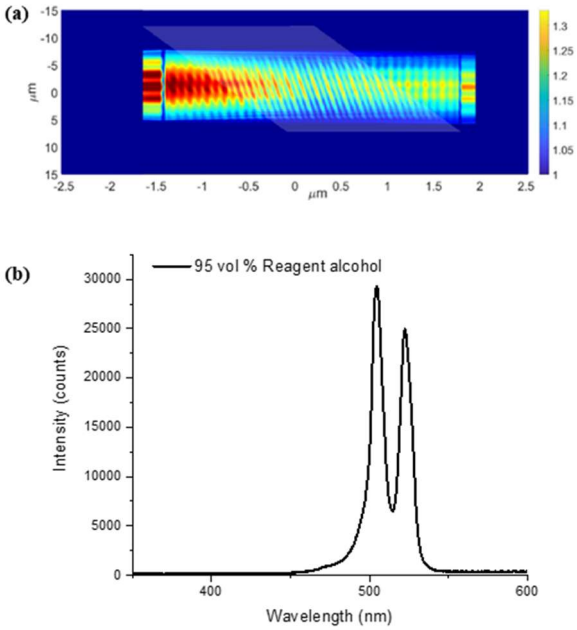


Figure 7. (a) Results of Bragg grating simulation. The laser beams only travel in the x direction. The pointing vector is visualised with a MATLAB built-in colour map called 'jet', which indicates the relative intensity of the energy flux. The intensity is highest at the radiation source, as the exposure flux is summed across the time steps. The periodic dark and bright fringes indicate the position of nodes and antinodes of the resulting standing wave from the counterpropagating laser beams respectively.
(b) An example of non-monochromatic spectrum (95 vol% reagent alcohol solution) due to the non-uniform intensity distribution as showed in (a).

The results from MATLAB simulation is shown in figure 7(a); this is a frame that visualises the output data towards the end of the FDTD time step loop. It depicts a clear interference pattern, with periodic dark and bright fringes that represent the nodes and antinodes of the standing wave respectively. The matrix that is being visualised is called the Poynting vector and is calculated by *Eqn.(9)*. It represents the directional energy flux of EM waves; in this instance, we have a resultant wave, therefore this quantity represents the power intensity that each unit area of polymer material is exposed to.

$$S = E \times H \qquad (9)$$

A close inspection of figure 7(a) shows that the distribution of intensity is non-uniform. This is shown by the unevenness in the fringes. This is a realistic reflection of the working condition of the laser which was found to produce patchy Bragg gratings, which may cause replayed wavelength to be less monochromatic due to different intensity of exposure during fabrication, as shown in figure 7(b), where two peaks are detected at certain detector orientations. However, as demonstrated in the rest of the characterisation results figure 5(a)-(d) showed that, for concentrations lower than 60 vol% in particular, the colorimetric response is sufficiently distinct for the sensor to act as an indicator for analyte concentrations. This suggests that the refractive index profile is successfully established even with an uneven laser exposure energy. This low level of sensitivity to spatial energy uniformity could suggest that the thickness of polymer medium plays a more important role in ensuring the selectivity of the Bragg grating; the higher number of holographic fringes as a result of thicker polymer film, as compared with previous works, may have resulted in ensuring a monochromatic and selective replay.

Conclusion

A double polymerisation process is used to fabricate a reusable, reversible, continuous holographic alcohol sensor. A range of different alcohols are tested, and the replay wavelength is used to characterise the internal refractive index profile due to volumetric changes upon exposure to analyte solution. The effect of different alcohols and different concentrations is investigated; the system displays linear diffraction wavelength response up to 40 vol% for all alcohol that are fully miscible in water. The visual readout can be used to identify the concentration of a known alcohol in DI water solution with a resolution of 10 vol%, with no sample treatment, sensor recalibration, and technical expertise required. In the work described in this report, a pHEMA based sensor with crosslinkers EDMA and GDM has exhibited ability to both quantify a known alcohol in water solution, as well as identify relative quantity of different alcohols in an unknown mixture.

The simulation results demonstrate that thicker polymer medium with greater number of holographic fringes increase the selectivity of the sensor, ensuring monochromatic replay and counterbalancing the effect of a non-uniform fabrication laser beam.

It is clear that this holographic sensor can be used in industrial beverage manufacturing processes to improve the monitoring of alcohol concentrations with further development. This is a label-free measurement pathway that serves as an optical transducer for colorimetric readouts, and does not involve nanoparticles, compared to current available holographic sensors; therefore, it stands out as a promising candidate for further development in the medical and biotechnological fields. Recommended further studies in the following direction are suggested: firstly, the conditions could be optimised in order to fabricate sensors to differ higher concentrations of alcohols with shorter response time. Secondly, diffraction efficiency could be calculated to compare the sensor brightness quantitively with formulation of different ratios. Thirdly, integration time optimisation at different alcohol concentration ranges, to obtain peaks with their maximum intensity reached, it would provide more reliable results for practical applications. Finally, using a light source with normally distributed intensity in the diffracted light region to reduce the influence of incident light intensity on the brightness of diffracted light and could compare the effect of concentration.

Acknowledgements
We would like to give special thanks to Jeff Blyth, Yunuen Montelongo, Samuel Davies for their help throughout this research project.

Supplementary Information
Research project- Alcohol senstive holographic sensor SI file.docx

References
Buchi.com. (2019). KjelFlex K-360 | buchi.com. [online] Available at: https://www.buchi.com/en/products/kjeldahl/kjelflex-k-360 [Accessed 10 Dec. 2019].
Anton-paar.com. (2019). Alcolyzer Spirits M/ME - Alcohol Analysis System :: Anton-Paar.com. [online] Anton Paar. Available at: https://www.anton-paar.com/uk-en/products/details/alcolyzer-spirits-mme-alcohol-analysis-system/ [Accessed 10 Dec. 2019].
Yetisen, A., Butt, H., da Cruz Vasconcellos, F., Montelongo, Y., Davidson, C., Blyth, J., Chan, L., Carmody, J., Vignolini, S., Steiner, U., Baumberg, J., Wilkinson, T. and Lowe, C. (2014). Light-Directed Writing of Chemically Tunable Narrow-Band Holographic Sensors. Advanced Optical Materials, 2(3), pp.250-254.
Mayes, A., Blyth, J., Kyröläinen-Reay, M., Millington, R. and Lowe, C. (1999). A Holographic Alcohol Sensor. Analytical Chemistry, 71(16), pp.3390-3396.
Oceanoptics.com. (2019). OceanView 1.6.7 | Ocean Optics. [online] Available at: https://oceanoptics.com/product/oceanview/?utm_campaign=D.+DSA+UK&utm_term=&utm_medium=ppc&utm_source=adwords&hsa_kw=&hsa_mt=b&hsa_grp=74937250703&hsa_tgt=dsa-40824091700&hsa_net=adwords&hsa_cam=2063190927&hsa_ver=3&hsa_acc=5588456569&hsa_src=g&hsa_ad=364112485160&gclid=Cj0KCQjwuNbsBRC-ARIsAAzITue7Z_syL0zpQPgsLzsLQ0VkjZ-HSY_kljLZXIRPgL5XtX3q41hoYtgaApRyEALw_wcB [Accessed 10 Dec. 2019].
K.S. Yee, “Numerical solution of initial boundary value problems involving Maxwell’s equations in iso-tropic media,” IEEE Transactions on Antennas and Propagation, vol. 14, pp. 302-307, 1966.
Imamoglu, T. and Yagci, Y. (2019). Photocuring of acrylate oligomers containing pentabromo benzylacrylate as flame-retarding monomer and properties of cured films. [online] Available at: https://onlinelibrary.wiley.com/doi/pdf/10.1002/app.2284.
Yuk, H., Zhang, T., Lin, S., Parada, G. and Zhao, X. (2015). Tough bonding of hydrogels to diverse non-porous surfaces. Nature Materials, 15(2), pp.190-196.
Benton, S. and Bove, V. (2008). Holographic imaging. Hoboken: Wiley-Interscience, pp.174, Ch. 16: In-line “Denisyuk” Reflection Holograms.
Yetisen, A., Qasim, M., Nosheen, S., Wilkinson, T. and Lowe, C. (2014). Pulsed laser writing of holographic nanosensors. Journal of Materials Chemistry C, 2(18), p.3569.
Thomas, R., Shoemaker, C. and Eriks, K. (1966). The molecular and crystal structure of dimethyl sulfoxide, (H3C)2SO. *Acta Crystallographica*, 21(1), pp.12-20.
Ballal, D. and Chapman, W. (2013). Hydrophobic and hydrophilic interactions in aqueous mixtures of alcohols at a hydrophobic surface. *The Journal of Chemical Physics*, 139(11), p.114706.
Sigmaaldrich.com. (2019). *Applications: Free Radical Initiators*. [ebook] Aldrich Polymer Products Applicaton & Reference Information, pp.12,16. Available at: https://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Aldrich/General_Information/photoinitiators.pdf [Accessed 11 Dec. 2019].
Marshall, A. J., et al. (2003). "pH-Sensitive Holographic Sensors." Analytical Chemistry **75**(17): 4423-4431.

Assessment of solvent performance in CCS process using molecular-based thermodynamic models

Yu Zhang

Dept. of Chemical Engineering, Imperial College London

London SW7 2AZ, U.K

Abstract

Solvent selection is crucial in the design of chemisorption-based carbon capture process. Due to intrinsic links between solvent molecules and process conditions, studies of the impacts of solvent structure on the process model must be conducted before implementing the solvent selection and process optimisation framework. Two series of parametric study on monoethanolamine (MEA) and 3-Aminopropanol were thus carried out to investigate the performance of these two solvents at various operating conditions. An equilibrium model is formulated to model a steady state absorber in the post-combustion carbon capture process. The absorber unit was implemented in gPROMS model-builder 5.1.1, with the use of state-of-art SAFT-γ-Mie EoS. Lower temperature of solvent feed was found to have higher CO_2 loading and less solvent loss but would require more energy input. Lower solvent concentration is favoured in terms of CO_2 loading and solvent loss but would increase the required liquid circulation rate, and thus lead to higher capital cost.

Introduction and Background

Climate change, especially global warming, is one of the most important issues in this century. In response to the threat of global warming, Paris agreement has set a target of limiting global temperature rise well below 2 degree Celsius above pre-industrial levels while pursuing efforts to limit it to 1.5°C. To reach this target, CO_2 emissions must be reduced to zero by 2050 (European Commission, n.d). Fossil fuel power plants are the major point-source of Carbon dioxide emission. Among the several types of carbon capture technologies, absorption based post-combustion capture is one of the most mature and viable method because of its 'end-of-pipe' characteristic.

Both physical and chemical absorption are broadly used in industries for the removal of CO_2 from gas streams. Comparing to physical absorption, amine-based chemical absorption has the advantage of effective removal of CO_2 in low concentration regime (Rao and Rubin, 2002). Solvent selection is crucial in the design of chemisorption-based carbon capture process. Three main types of amines, namely the primary, secondary and tertiary amines, are usually used as solvents for chemical absorption. The primary amines, such as monoethanolamine (MEA) has high reactivity, and therefore are most suitable to work at low CO_2 partial pressure. However, the large energy demand for the regeneration of this kind of amines becomes the main obstacle in the large-scale deployment

of primary amine-based post combustion capture process. According to the U.S Department of Energy (DOE)/National Energy Technology Laboratory (NETL), the deployment of current MEA based chemical absorption process in a coal-fired power plant would decrease the plant efficiency by 30% (DOE/NETL,2013; Li et al., 2016). Secondary and tertiary amines require less energy for regeneration because of weaker association bonds(Laddha and Danckwerts, 1981;Blauwhoff et al., 1984; Barth et al., 1984). However, their low reactivity with CO_2 makes them unsuitable for the process containing relative dilute CO_2. Solvent degradation is another main concern with commercial implementation of the post-combustion technology as it not only leads to potential loss of solvent but also brings health and environmental impacts. There are two main types of degradations, which are thermal degradation and oxidative degradation. Thermal degradation usually occurs in the stripper column because of the high operating temperature, and oxidative degradation usually takes place in the absorber column, where the solvent is in contact with oxygen present in the flue gas (Flø et al,.2017).

In attempt to advance the performance of chemisorption CCS, many experimental studies have been undertaken to identify potential candidate solvents that would outperform the current benchmark solvent, aqueous solutions of monoethanolamine (MEA) (Papadopoulos, 2016). With the advance in computer-aided molecular design method and the development of simplified thermodynamic properties model, CAMD has become the traditional key processes of identifying potential solvents. The use of group contribution method (GC) also allows unrestricted prediction of compound properties. In other words, it can predict the fluid properties, even for compounds that have never been synthesised before. Such molecular design method can only be applied to the case with given knowledge of process conditions and thus, the ultimate objective of optimizing both molecular and process design is achieved by addressing the two problems sequentially (Eden et al., 2004). However, the choice of processing materials, solvent in the case of CCS, is tightly linked to decisions on process topology and process conditions. Therefore, a CAMD method may lead to poor performance as it limits the design space and neglects the intrinsic links between molecules and process (Adjiman, Galindo and Jackson, 2014). Given the interdependent nature between solvent and process design, recent studies have given more focus to the process-wide computer-aided molecular and process design (CAMPD), which aims to identify optimum solvents and their associated optimal process conditions simultaneously.

Several types of approaches have since then been proposed for solving CAMPD. The majorities of existing approaches fall into the of decomposition and screening approach, in which the MINLP problem that arises from a CAMPD is solved by decomposing the original problem into a primal problem and master problem. One of the earliest decomposition approaches was proposed by Buxton et al. The approach considered a separation problem in which an optimum solvent blend is to be designed. Potential solvent blends are first screened with a series of feasibility and properties test. The mixture properties for each solvent blend are then solved as an NLP to identify the optimal blend compositions for each component. Recently, Gopinath et al. (2016) proposed an outer approximation algorithm with physical domain reduction to address a CO_2 physisorption process design problem. With the incorporation of a series of primal feasibility tests to enhance the numerical robustness of the design algorithm, the results showed that decomposition-based approach can be used effectively in solving CAMPD. However, considering all the benefits that chemisorption brings to the carbon capture processes, it would be instrumental to develop a reliable and predictive model of

chemisorption-based carbon capture processes using a similar CAMPD design algorithm.

Since solvent molecules and process conditions are interrelated, studies of the impacts of solvent structure on the process model are needed to gain a better understanding of the process. Brand et al. (2016) investigated the performance of the rated-based absorber model using aqueous MEA with the integration of SAFT-VR SW EoS. The study has provided useful insights into the connections between model behaviour and process and operating conditions. However, the SW EoS are known to have limitations in providing a simultaneous description of the fluid-phase behaviour and thermodynamic derivative properties due to the simplified representation of intermolecular forces. (Papaioannou et al.) A recent major advance to overcome such issues in SAFT is the development of contribution method based thermodynamic approach SAFT-γ-Mie EoS. In this model, the Mie potential is a generalised form of Lennard-Jones potential, which gives improved prediction of caloric properties, providing a good description of vapour-liquid equilibria (Lafitte and Apostolakou et al.,2013; Lafitte et al., 2006;Dufal et al., 2015a; Dufal et al., 2015b).

The objective of this research is to develop an amine-based carbon dioxide absorber model using the state-of-art SAFT γ-Mie EoS. In addition, parametric studies on 3-Aminopropanol and MEA are carried out in order to gain a better understanding of solvent performance with relation to operating conditions. and the effects of solvent concentration on CO_2 loading, effluent temperature and solvent loss.

Methodology

1.Thermodynamic model

The above mentioned SAFT-γ-Mie approach was implemented in gPROMS model-builder as a foreign object for the modelling of amine-based absorber column.

2.The process

Figure 1. shows the flowsheet of the amine based post-combustion carbon capture process considered in this study. The process involves a counter-current absorber and a counter-current desorber, with a liquid feed going into the absorber at the top stage and a vapour feed from the bottom stage. The rich loading solvent stream from the absorber is then fed to the top of the desorber for CO_2 stripping and solvent regeneration. As a recycle stream, the lean solvent from the desorber containing regenerated solvent is then fed back to the top of the absorber.

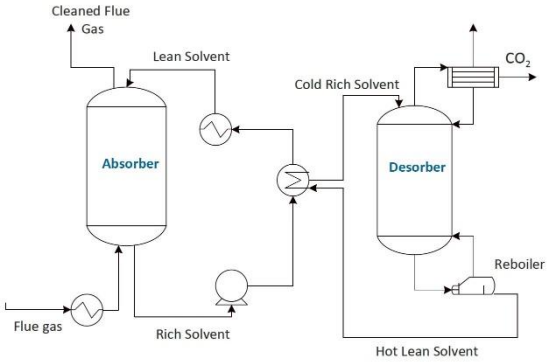


Figure 1. Schematic diagram of amine-based post-combustion carbon capture process

3. Absorber model

The absorber unit in the post-combustion carbon capture process can be modelled by two approaches, namely the rate-based and the equilibrium-based modelling. In constructing the rate-based model, the two-film theory has been mostly used to describe the behaviour at the gas-liquid interface (Flø, 2015).

In this context, the absorber was modelled using an equilibrium model. In this model, phase and chemical equilibrium is assumed to be achieved at each stage, and everywhere on the stage.

4. Model equations:

It is assumed that the pressure drop across the column is negligible. Therefore, the pressure of the liquid and vapour stream can be expressed as:

$$P_{j-1}^{L} = P_{j}^{L} \quad (1)$$

$$P_{j+1}^{V} = P_{j}^{V} \quad (2)$$

where P_j^L and P_j^V are the pressure of the liquid and vapour stream leaving stage j respectively.

In addition, the absorber is modelled at steady state. Therefore, there is no accumulation of mass or energy in the process system.

Mass balance:

The component mass balance for component i at stage j in liquid and vapour phases can be expressed though the following equations:

$$L_{j-1}x_{i,j-1} = L_j x_{i,j} \quad (3)$$

$$V_{j+1}y_{i,j+1} = V_j y_{i,j} \quad (4)$$

where L_j and V_j are the total liquid and vapour molar flowrate leaving stage j respectively, and $x_{i,j}$ and $y_{i,j}$ correspond to the mole fractions of component i in the liquid and vapour stream leaving stage j.

The sum of the mole fractions of the components in the stream leaving stage j must equal to unity.

$$\sum_i x_{i,j} = 1 \quad (5)$$

and

$$\sum_i y_{i,j} = 1 \quad (6)$$

Heat balance:

Since the absorber was assumed to operate adiabatically, the energy balance for the absorber column can be written as

$$L_{j-1}H_{j-1}^{L} + V_{j+1}H_{j+1}^{V} = L_j H_j^L + V_j H_j^V \quad (7)$$

where H_j^L and H_j^V are the enthalpies of the liquid stream and vapour stream leaving stage j respectively. H_j^L and H_j^V are obtained from SAFT-γ-Mie as functions of temperature T_j^L and T_j^V, pressure, and composition of each phase $x_{i,j}$ and $y_{i,j}$.

Equilibrium relations:

Both chemical and physical equilibrium were assumed to be achieved everywhere on the stage. Therefore, the temperature, pressure and chemical potential of the liquid phase must be equal to that of the vapour phase:

$$T_j^L = T_j^V \quad (8)$$

$$P_j^L = P_j^V \quad (9)$$

$$\mu_j^L = \mu_j^V \quad (10)$$

where μ_j^L and μ_j^V are the chemical potentials, defined as functions of temperature, pressure and composition in SAFT-γ-Mie EoS, of the liquid and vapour stream leaving stage j respectively.

The composition $x_{i,j}$ and $y_{i,j}$ were calculated through TPFlash, an equilibrium flash method in Multiflash, gPROMS, with equilibrium temperature and pressure, and the overall mixture compositions as method inputs.

5. Heat of absorption

Heat of absorption, same as enthalpy of absorption, is an important factor that affects the energy requirement of the CO_2 capture process. In order to the assess the model prediction accuracy, the predicted values of heat of absorption of CO_2 in 30wt% MEA by SAFT-γ-Mie are compared with a set experimental data (Arcis et al., 2011). The experiment measured the enthalpy change of absorption at constant temperature and pressure as a function of CO_2 loading, α_{CO_2}(mol CO_2/mol amine), which in this experiment was defined by the total flowrate of CO_2

divided by the total flowrate of amine. As shown Figure 2., SAFT- γ-Mie EoS slightly underestimates the enthalpy of absorption of CO_2 by about 10kJ/mol at 322.5 K, 0.51MPa.

6. Absorber model validation

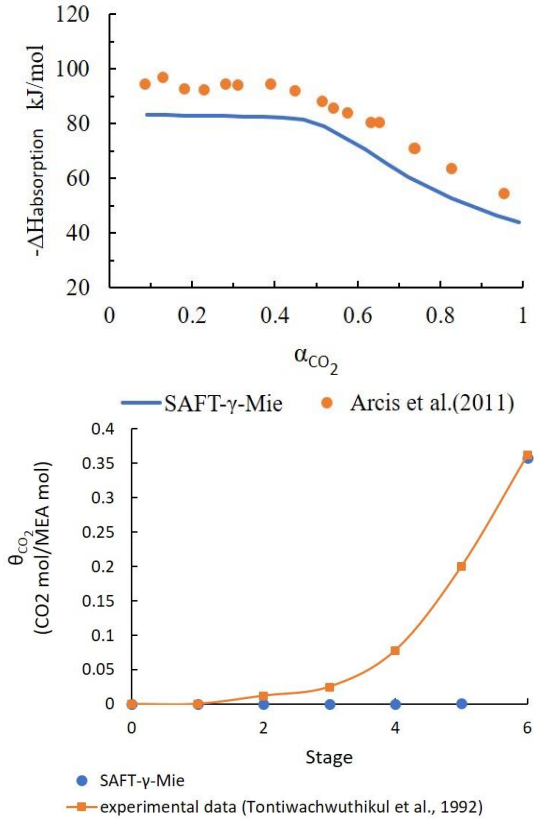


Figure 3. Comparison of CO_2 loading profile predicted by my model and that from literature

CO_2 loading profile and liquid stream temperature profile from model prediction were compared to literature (Tontiwachwuthikul et al., 1992)) Run T13 using same operating conditions. $\theta_{CO_2}$is defined by the molar flowrate of CO_2 in the liquid phase divided by the molar flowrate of amine in the liquid phase. As seen in Figure 3, the model achieves same value of CO_2 loading at the bottom stage as the pilot plant data. However, without taking mass transfer limitations into account, the model overpredicts the rate of absorption.

Figure 4 shows that the model overestimates the temperature of the liquid leaving from the bottom stage. Since the energy released from reaction between CO_2 and MEA solution is assumed to be the only source of heat to the absorber, and the model underestimated the total energy released from reaction in this scenario, one can draw the conclusion that the overestimation of rich solvent temperature may be a resulted from the underestimation of the heat capacity of liquid mixture.

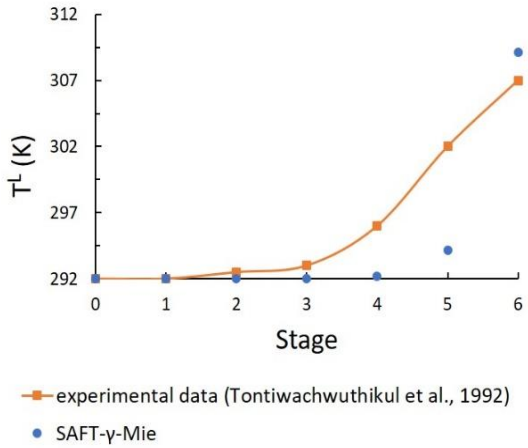


Figure 4. Comparison of Temperature profile predicted by my model and that from literature

Results and discussion:

Case 1: Vary Liquid inlet temperature

The temperature of solvent feed to the absorber column was varied from 276K to 330K to investigate its effects on CO_2 loading, effluent temperature and solvent loss. As shown in Figure 5, MEA and 3-Aminopropanol have very similar values of maximum CO_2 loading, and these maximum values occur also at similar solvent feed temperatures. As temperature of solvent feed stream increases, the CO_2 loading at bottom stage for both MEA and 3-Aminopropanol decrease. i.e less CO_2 have been absorbed by the solvents. This trend can be directly related to the exothermic characteristic of the absorption reaction. When CO_2 is absorbed by a solvent, energy is released through the exothermic reaction between the two species, subsequently leading to an increase in the

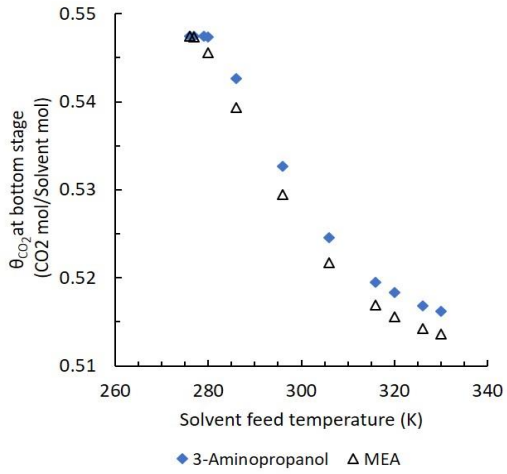


Figure 5. Relation between bottom stage CO_2 loading and the temperature of feed solvent

temperature of CO_2-solvent mixture. Since a relatively higher temperature would shift the reaction equilibrium of an exothermic reaction towards the reactant, one would expect a decrease in the CO_2 loading as the mixture temperature increases. It can also be noticed that CO_2 loading for MEA is less than that for 3-Aminopropanol at any temperature between 280K and 330K. Therefore, low solvent temperature and 3-Aminopropanol are preferred in terms of CO_2 loading.

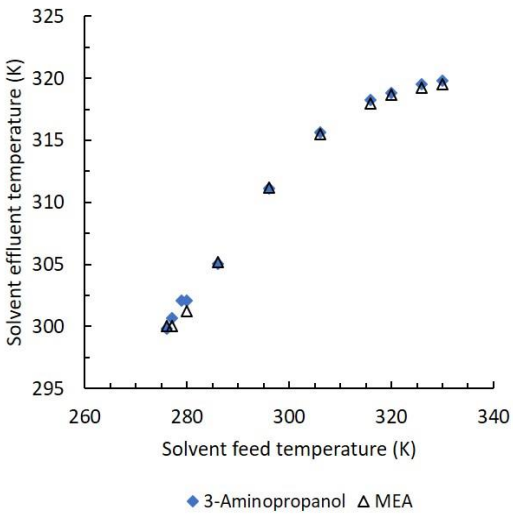


Figure 6. Temperature relation between solvent feed and effluent

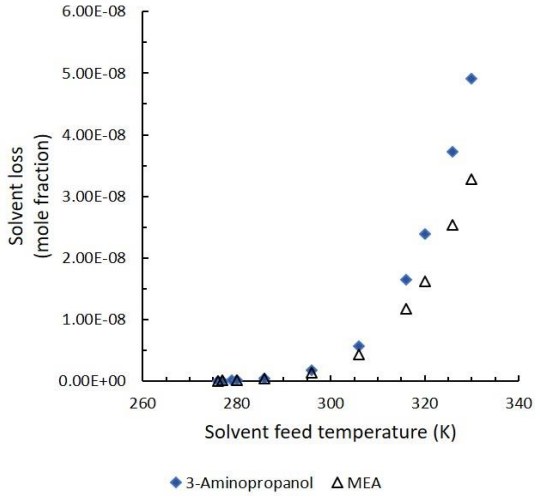


Figure 7. Solvent loss against solvent feed temperature

In the absorber model, vaporisation is the only source of solvent loss that has been considered because the model builder has limited ability to simulate thermal degradation. Further assumption is made that there is no oxygen present in the flue gas stream to cause oxidative degradation. Therefore, degree of vaporisation increases when the solvent feed temperature increases, and thus results in more solvent loss. As seen in Figure 7, the value of solvent loss increases exponentially with increase in temperature. However, it is surprising to see 3-Aminopropanol has more significant solvent loss than MEA does at higher solvent feed temperature because pure 3-Aminopropanol is less volatile than pure MEA. Therefore, Figure

8. VLE diagram of MEA and 3-Aminopropanol was plotted to investigate the causes behind this

observation. It can be seen from the plot that a reverse trend in bubble point of aqueous MEA and aqueous 3-Aminopropanol occurs when the mole fraction of amines in aqueous solution is reduced from 1 to about 0.75, indicating 3-Aminopropanol becomes more volatile than MEA when their concentration is lower than 0.75 by mole fraction. The increase in the relative volatility of aqueous 3-Aminopanol to

aqueous MEA is due to the presence of extra hydrophobic CH_2 group in 3-Aminopropanol

Table 1. Inputs to absorber column for case 1

	Flue gas feed	**Solvent feed 3-Aminopropanol**	**Solvent feed MEA**
Flowrate (mol/s)	0.14	Vary with amine concentration	Vary with amine concentration
Components	N_2, CO_2, H_2O	H_2O 3-Aminopropanol	H_2O MEA
Amine concentration (mole fraction)	0	Vary from 0.02 to 0.2	Vary from 0.02 to 0.2
P_{CO_2} (MPa)	0.0153	0	0
Pressure (Mpa)	0.1	0.1	0.1
Temperature (K)	292	292	292

molecule that contributes to the non-polarity, and thus weakens the interaction with water. In brief, lower solvent feed temperature is required to reduce solvent loss. In addition, the attribution of interaction between solvent and water molecules can also greatly affects solvent loss. In general, molecules that exhibit stronger intermolecular interactions are preferred in term of solvent loss reduction.

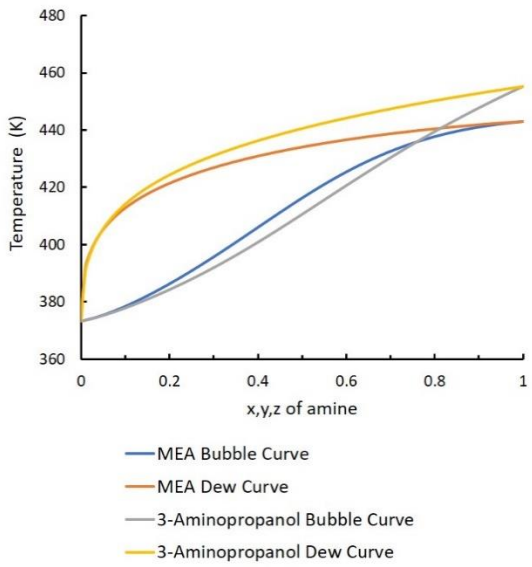


Figure 8. VLE diagram of aqueous amine solution at 0.1MPa predicted by SAFT- γ-Mie

Case 2: Vary solvent concentration

In addition to solvent inlet temperature, inlet solvent concentration has been found to have significant impacts on CO_2 loading and solvent loss. In this case, the recovery of CO_2 was fixed to be 0.1 to draw a comparable result of performance of solvents at different concentrations.

Figure 9 shows that CO_2 loading decreases as solvent concentration increases, indicating a relatively higher concentration of these amine solutions would require larger amount of amine presented in the solvent feed in order to absorb same amount of CO_2 than the relatively less concentrated amine solution does.

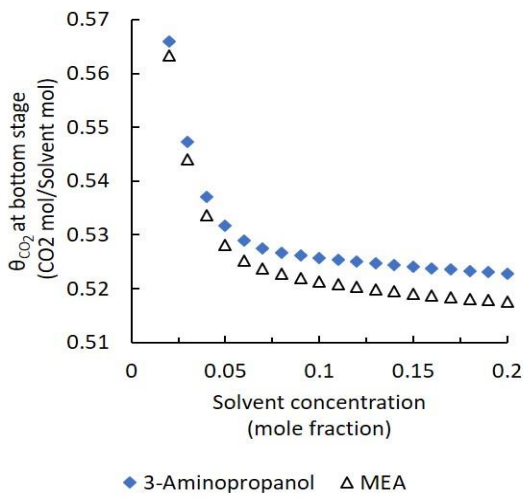


Figure 9. CPO_2 loading against solvent concentration

This finding also agrees with the solubility data. Figure 10. reveals the relationship between CO_2 loading and the concentration of absorbent at different P_{CO_2}. In the absorber model, P_{CO_2} in the flue gas stream is usually around 10^4 Pa, which corresponds to the top region of the solubility graph. In this region, the higher the

Table 2. Inputs to absorber column for case 2

	Flue gas feed	Solvent feed 3-Aminopropanol	Solvent feed MEA
Flowrate (mol/s)	0.14	1.0	1.0
Components	N_2, CO_2, H_2O	H_2O 3-Aminopropanol	H_2O MEA
Amine concentration (mole fraction)	0	0.0385	0.0385
P_{CO_2} (MPa)	0.0153	0	0
Pressure (MPa)	0.1	0.1	0.1
Temperature (K)	292	Vary from 276 to 330	Vary from 276 to 330

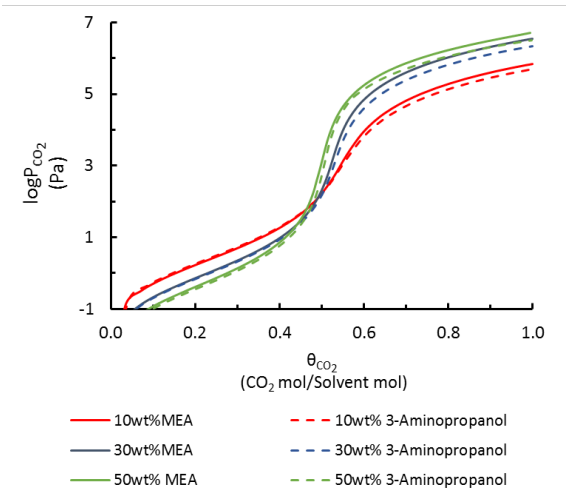


Figure 10. Solubility of CO_2 in aqueous amine solutions

concentration of MEA and 3-Aminopropanol solvents, the lower the resulted CO_2 loading.

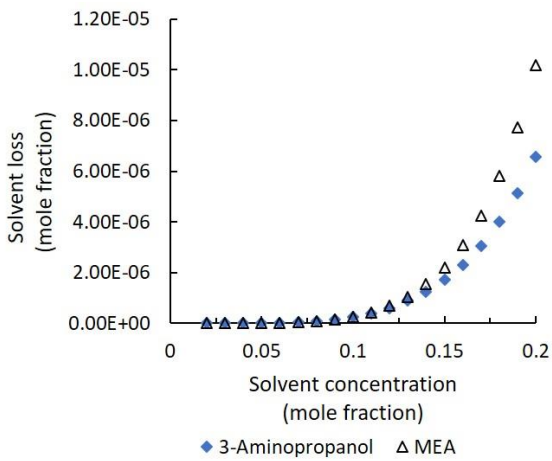


Figure 11. Solvent loss against solvent concentration

Figure 11 was plotted for both test solvents in order to get a better understanding on the effects of solvent concentration on solvent loss. As can be seen from the graph, solvent loss shows an increasing exponential trend as the inlet solvent concentration goes up. Taking MEA as an example, the point at which solvent mole fraction equal to 0.1 corresponds to 30wt%. Since solvent loss becomes significant at any concentrations higher than this point, a concentration lower than 30wt% is preferable in term of maintaining relatively small solvent loss.

The more significant solvent loss for MEA at high solvent concentration as revealed in Figure 11 is due to the higher temperature of the cleaned gas stream, T_1^V, of MEA. At relative higher concentration, temperature has greater impact on solvent loss than the interaction between solvent and water molecules does.

Conclusion

Two series of parametric study on MEA and 3-Aminopropanol were conducted to investigate the performance of MEA and 3-Aminopropanol at various operating conditions for absorber column in an amine based post-combustion carbon capture process. It was found that a lower temperature of solvent feed results in achieve a higher CO_2 loading and less solvent loss but would require more cooling duty for the cooler and more sensible heat in the desorber unit. Lower solvent concentration is favoured in terms of loading and solvent loss. However, the resulted solvent circulation rate would be higher, and therefore would result in higher capital cost for a larger size absorber column; 3-Aminopropanol has better performance than MEA in terms of CO_2 loading.

Outlook

In the work, a process model has been built for a stand-alone absorber column. There is future scope of performing the study the performance of solvents in relation to operating conditions in the desorber. By doing this, one would be able to perform quantitative analysis on the overall carbon capture process.

Since the result of this study shows that optimizing one process variable would result in both pros and cons, a multi-objective global assessment on the performance of three types of amine on the complete carbon capture process are required.

Acknowledgement

I would like to acknowledge Lauren Lee for her patience, guidance and support throughout this project.

Nomenclature

i	Component index
j	Stage Index
P_j^L	pressure of the liquid leaving stage j
P_j^V	pressure of the vapour leaving stage j
x	Mole fraction of component in liquid stream
y	Mole fraction of component in vapour stream
T_j^L	Liquid stream temperature at stage j (K)
T_j^V	Vapour stream temperature at stage j (K)
P_j^L	Liquid stream pressure at stage j temperature (Pa)
P_j^V	Liquid stream pressure at stage j temperature (Pa)
μ_j^L	Chemical potential of liquid stream at stage j (J/mol)
μ_j^V	Chemical potential of vapour stream at stage j (J/mol)

References

[1]Adjiman, C. S., Galindo, A., and Jackson, G. (2014). "Molecules matter: The expanding envelope of process design". In Eden, M. R., Siirola, J. D., and Towler, G. P., editors, Proceedings of the 8th International Conference on Foundations of Computer-Aided Process Design, volume 34 of Computer Aided Chemical Engineering, pages 55–64. Elsevier

[2]Arcis, H., et al. (2011). "Enthalpy of Solution of Carbon Dioxide in Aqueous Solutions of Monoethanolamine at Temperatures of 322.5 K and 372.9 K and Pressures up to 5 MPa." Journal of Chemical & Engineering Data 56(8): 3351-3362.

[3]Brand, C. V., et al. (2016). "On the use of molecular-based thermodynamic models to assess the performance of solvents for CO2 capture processes: monoethanolamine solutions." Faraday Discussions 192(0): 337-390.

[4]Eden, M. R., Jørgensen, S. B., Gani, R., and El-Halwagi, M. M. (2004). "A novel framework for simultaneous separation process and product design". Chemical Engineering and Processing: Process Intensification, 43:595–608.

[5]European Commission. (n.d). "Paris Agreement." Retrieved 02 November, 2019, from https://ec.europa.eu/clima/policies/international/negotiations/paris_en.

[6]Flø, N. E., et al. (2017). "Results from MEA Degradation and Reclaiming Processes at the CO2 Technology Centre Mongstad." Energy Procedia 114: 1307-1324.

[7]Gopinath, S., et al. (2016). "Outer approximation algorithm with physical domain reduction for computer-aided molecular and separation process design." AIChE Journal 62(9): 3484-3504.

Laddha, S. S. and P. V. Danckwerts (1981). "Reaction of CO2 with ethanolamines: kinetics from gas-absorption." Chemical Engineering Science 36(3): 479-482.

[8]Lafitte, T., et al. (2013). "Accurate statistical associating fluid theory for chain molecules formed from Mie segments." The Journal of Chemical Physics 139(15): 154504.

[9]Li, Z., et al. (2016). "Verification of a solvent optimization approach for postcombustion CO2 capture using commercial alkanolamines." International Journal of Greenhouse Gas Control 44: 59-65.

[10]N. E. Flø, "Post-combustion absorption-based CO2 capture: modeling, validation and analysis of

process dynamics," PhD-dissertation, Chemical Engineering, NTNU, Trondheim, Norway, 2015
[11]Nguyen, T., et al. (2011). "Volatility of aqueous amines in CO2 capture." Energy Procedia 4: 1624-1630.
[12]Papadopoulos, A., et al. (2016). "Computer-aided molecular design and selection of CO2 capture solvents based on thermodynamics, reactivity and sustainability." Molecular Systems Design & Engineering 1: 313.
[13]Rao, A. B. and E. S. Rubin (2002). "A Technical, Economic, and Environmental Assessment of Amine-Based CO2 Capture Technology for Power Plant Greenhouse Gas Control." Environmental Science & Technology 36(20): 4467-4475.
[14]Tontiwachwuthikul, P., et al. (1992). "CO2 absorption by NaOH, monoethanolamine and 2-amino-2-methyl-1-propanol solutions in a packed column." Chemical Engineering Science 47(2): 381-390.

Optimisation and Evaluation of CHP Engine Operational Strategies within Sainsbury's Stores

Toby Daniels and Bhanu P. Pillay

Department of Chemical Engineering, Imperial College London, U.K.

Abstract

This paper reviews the widening of 250kWe CHP engine operational constraints recommended by the engine suppliers and covers the MINLP optimisation strategy involved in modelling them. Using this model, financial and environmental multi objective optimisation is applied to 15 Sainsbury's facilities datasets, giving estimates of potential savings of up to 31% and a consistent reduction in carbon emissions relative to baseline strategies. Estimated savings associated with the introduction of a 250kWe CHP engine is found to increase linearly with store size, with smaller stores, however, benefitting more greatly from the designed optimiser. Sensitivity analysis applied to the model demonstrates additional saving of up to 10% corresponding to increases in electricity export price prediction accuracy, with rising gas and electricity prices associated with a decrease and increase in savings, respectively. Using a variety of data sets, modelled engine maintenance costs and a realistic engine operational domain, results from this optimisation strategy ultimately allow for possible development and evaluation of potential future Sainsbury's investment into CHP technology and adoption of various operational strategies.

Introduction

Wellesbourne Sainsbury's

A mid-sized Sainsbury's store in Wellesbourne, Warwick is used as a case study for the investigation of the 250kWe engine optimisation. This engine having been installed on the Wellesbourne store site in junction with access to maintenance and engine specific data allows for the financial and environmental modelling in this report to be carried out. To lower the store's reliance on the grid to satisfy heating and electricity demand, the engine installed (250kWe) is oversized, and thus, the optimal engine setpoint is primarily dictated by electricity import and export prices. Both real and forecasted electricity price data is obtained alongside Wellesbourne store data, which is then incorporated into the analytics of other stores. The Wellesbourne Sainsbury's ultimately gives us a platform for preliminary results of optimisation constraint implementation, eluding to potential wider applicability to other stores.

Potential savings

The use of a combined heat and power system allows for both sale of generated electricity back to the grid and produced heat being used for store space heating if necessary. CHP engine optimisation allows for dynamic control of the engine setpoint, enabling a higher output when electricity prices are high and a lower output when prices are low, alongside optimally satisfying the store's heat and power needs. This provides an avenue for financial savings as well as outlining operational strategies to reduce carbon emissions.

Sainsbury's Partnership Work and Literature Review

In 2010, Sainsbury's established a partnership with Imperial College London with the goal of pursuing research to reduce its carbon footprint from retail activities. The partnership is constantly reviewing technologies and innovative approaches to enhance business operations. Extensive analysis into optimal selection of various control strategies for the engine has been carried out (Yining, L, 2018), a load following, a business as usual strategy and an ON/OFF strategy. Linear engine efficiency relations have been formulated as well as fuel carbon indexes, allowing for control strategy equations to be coded into Python. The identification of these parameters and the subsequent Python model developed provides a base for the strategies investigated in this report. It should be noted that an older version of the cost function is used for the optimisation model developed in this paper, but still provides valuable insight into savings associated with modelled constraints. Optimisation of allocation of a range of technologies

has also been investigated, including bio-methane and natural gas scenarios (Alvarado, D. C, 2016). Analysis of CHP only and CHP – ORC (Organic Rankine Cycle) systems have allowed for selection of these systems based on their associated carbon emissions and forecasted costs (Mago, P. J, 2010). The models here utilise either single or multi-objective optimisation, outlining operational strategies corresponding to various socio-environmental climates, a technique also used in this project.

Project Objectives and Overview

The objectives of this project are outlined through the following research questions:

-What CHP engine constraints confirmed by the engine supplier can be practically implemented into an optimisation model to increase savings and reduce environmental impact
-What are the benefits of widening the operational domain on the engine in this way?
-Which stores, and which store characteristics are most suitable for the implementation of the 250kWe CHP engine and the designed optimiser?
-How does the model adapt to minimise carbon emissions of the engine and what are the corresponding changes to savings?

Sequential completion of specific tasks is required to comprehensively answer the research questions. Initially, a preliminary review of additional savings from postulated constraints applied to the Wellesbourne store data provides key indications of which constraints to further investigate. Information then obtained from the engine supplier (Centrica) confirms which constraints lie within safe and practical operational limits, as well as generalised CHP maintenance and capital costs. This allows for the design of a realistic Pyomo optimisation model to apply these constraints in conjunction with the pre-existing cost function. After analysis of the Sainsbury's store data base, stores with compatible input data for the cost function are identified and the designed optimiser run for the associated data sets. Additional savings corresponding to the Centrica confirmed constraints are calculated for the highlighted stores, with pareto curves indicating the impact of low carbon strategies. Through data visualisation and analytics, store characteristics are then analysed against financial and environmental improvements, to help predict which stores benefit most from implementation of the 250kWe engine and the designed optimiser.

Methodology

Cost Function

Having already been developed as part of the Imperial-Sainsbury's partnership, the cost function used in this project calculates the cost associated with each 30-minute timestamp, additively incorporating a weighted carbon cost to be used for the environmental impact assessment. Previously gathered data sets are used as inputs for the function, including heat and electricity demand, predicted and actual electricity prices (Exelon, 2019), grid carbon intensity and the CHP power output, the latter corresponding to the setpoint of the engine. Fuel price and engine specific coefficients have been obtained through internal discussions with engine suppliers and retailers and are used to give an estimate of the carbon and financial costs. The linear fuel consumption and heat generation relationships are shown in (*Eq. 1 & 2)*.

$$fuel_{CHP} = 2.64 \times elec_{produced} + 76.6 \quad (1)$$

$$heat_{CHP} = 1.06 \times elec_{produced} + 48.7 \quad (2)$$

Costs and carbon outputs can then be calculated for the boiler and engine fuel in (*Eq. 3,4,5 & 6)*.

$$fuel_{cost} = fuel_{CHP} \times 2.54, \quad (3)$$

$$fuel_{carbon} = fuel_{CHP} \times 0.184, \quad (4)$$

$$boiler_{cost} = 2.54 \times (d_h - heat_{CHP}), \quad (5)$$

and $boiler_{carbon} = 0.184 \times (d_h - heat_{CHP}),$ (6)

where d_h denotes the heating demand for the store for the respective timestamp. A similar methodology is followed to calculate total electricity costs and grid carbon impact; the full function can be found in *Appendix A.* An annualised CHP engine maintenance cost of £15,000 was given by Sainsbury's engineers at Wellesbourne. As this corresponds to the operational strategy of the engine being at a fixed setpoint during business hours, the figure can be factored down to estimate the maintenance cost per hour of the engine being on. This value, £3.43 per hour, is implemented into the cost function to allow for a more accurate modelling of the CHP cost. Carbon released is calculated as a sum of the boiler, fuel, grid contributions and displacements, with the

final cost being an additive combination of the cost and weighted carbon sums as shown in *Eq. 7*.

$$Total\ Cost = Cost + Total\ Carbon \times Carbon_{factor}, \quad (7)$$

where $Carbon_{factor}$ varies arbitrarily depending on the value placed on reducing carbon emissions. Although it doesn't represent an actual cost and bares no physical relevance, it ultimately allows for a poreto curve to be plotted corresponding to a multi objective optimisation model. The cost function accounts for scenarios when excess electricity is produced and exported to the grid, and when the CHP engine doesn't satisfy energy demands and electricity must be imported. A predicted electricity export price algorithm already developed is used when determining the optimal engine setpoint. Using predicted export prices, the cost function is run for multiple engine setpoints in a specified range for each timestamp. The setpoint corresponding to the lowest cost is selected and plugged back into the cost function using the actual price data to output the final cost. It should be noted that to calculate final cost here, the carbon factor is set to 0.

Comparison strategies

To assess the impact of the optimiser outlined in this paper, it is compared to 2 base strategies, a fixed setpoint and load following case. Fixed setpoint operation at 162kWe (65%) during business hours corresponds to a business as usual (BAU) case and is the current engine strategy at the Wellesbourne case study site. The load following case is where the engine is operated at a setpoint corresponding to the electricity demand of the store at that time during business hours, irrespective of external market factors. With this strategy, if the power requirements of the store are higher than the maximum engine output, the setpoint is set to 100%. Likewise, if the requirements are lower than the minimum setpoint, output is set to the minimum value. The load following case allows for comparison with the most practical 'base' engine strategy if setpoint optimisation is not possible, and the BAU case gives a realistic comparison to the current Sainsbury's operational strategy for the 250kWe engine.

Store selection

Out of the 25 Sainsbury's buildings of which data is available through the partnership, 15 were selected to have compatible input data for the cost function to allow for optimisation. The building ID's investigated in this paper along with their locations are shown in *Table A* in *Appendix B,* 3 of these not being stores (2 distribution centres and 1 corporate office). The selection of multiple buildings with varying locations, demand profiles and sizes allow for profiling of building characteristics against both the application of the designed optimiser and the base CHP engine strategies.

Preliminary Data Analytics

When adding a possible engine setpoint of 0, allowing the engine to turn off, a profile of optimal engine behaviour for Wellesbourne is calculated with the cost function. The engine is seen to turn on around 7-8am, following the store load until around midday, at which it switches off. Around 4-5pm, the engine ramps up to full output (250kWe), taking advantage of high electricity export prices during these times before ramping down and turning off in the late evening. Despite the engine being off for a few individual timestamps throughout the day, the consistent 2-4 hour 'off' duration in the afternoon eludes to potential savings even if the number of ON/OFF switches per day is limited.

This paper investigates the implementation of certain constraints to widen the operational domain of the CHP engine. Lowering the minimum output of the engine, increasing the allowable operation time at a low engine output, and increasing the number of allowable daily ON/OFF switches are investigated. Preliminary data analytics is applied to a 6-week dataset (08/19 – 09/19) obtained from the Wellesbourne Sainsbury's store to indicate potential savings improvements corresponding to the mentioned constraints.

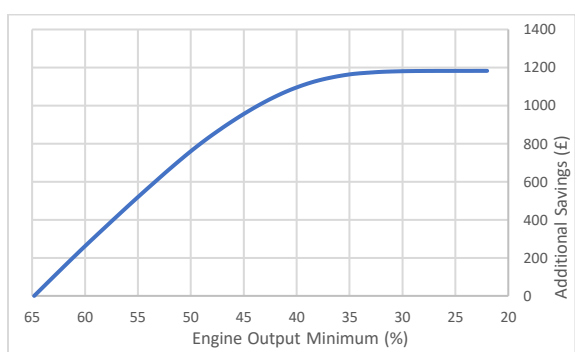


Figure 1: Savings increase against minimum engine output

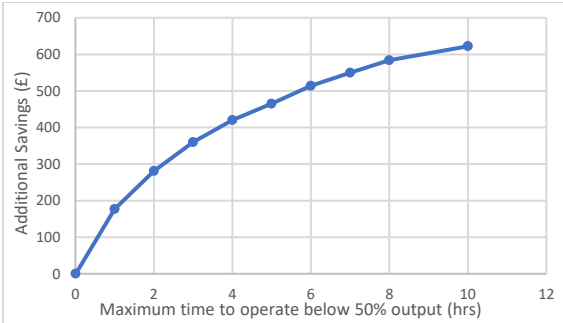


Figure 2: Savings increase against maximum time operating below 50% output

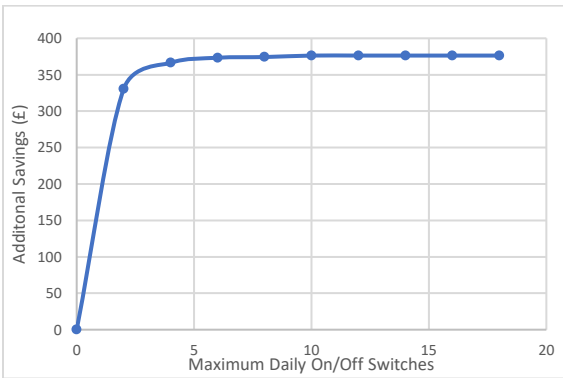


Figure 3: Savings increase against maximum daily On/Off switches

Fig. 1 demonstrates savings of £1200 across the 6 week dataset as the minimum output of the 250kWe engine is allowed to decrease to 25%, with the additional savings increasing linearly before plateuing at 30%. Furthermore, as expected, additional savings rise as the allowable daily number of hours operating below 50% increases to 10, as seen in *Fig. 2*. *Fig.3* shows that when increasing maximum daily ON/OFF switches, savings increase by £380, with savings plateuing after 4-6 switches. Additionally, carbon emissions decreases by up to 11% relative to the BAU case for Wellsbourne when increasing daily ON/OFF switches. Grid sourced electricity being modelled to be less carbon intesive than the CHP engine, allows for forecasted carbon emissions to decrease when the engine is allowed to turn off throughout the day. Optimising the Wellsbourne dataset for the investigated constraints demonstrate significant potential savings, increasing confidence in the relavence of exploring these constraints for the engine.

Centrica Recommendations

Upon request, the CHP engine supplier for Sainsbury's, Centrica, has detailed comprehensive engine operational limits. Specifically, corresponding to the constraints preliminarily investigated, Centrica recommends that:

-The minimum engine output is not to fall below 40%
-The engine can operate at a maximum of 3 hours per day below 50% output
-If the engine is to operate below 50% output for at least 1 hour in a day, the engine must then operate at 100% in the same day for at least 2 hours in order to remove engine soot build-up
-A maximum of 6 daily ON/OFF switches for the engine; turning on then off counts as 2 switches

Relating to the Wellesbourne dataset, Centrica recommended constraints can be seen to still result in notable savings, with 6 ON/OFF switches resulting in the maximum 'ON/OFF' savings being incurred. Once these practical constraints are known, the optimisation strategy can be applied to all relevant Sainsbury's building's datasets to review the benefits of the optimiser relative to the base strategies.

MINLP Model

Formulation Strategy

As engine constraints are given per day, the Pyomo model will optimise consecutive daily datasets, outputting a vector of 48 optimised costs (each corresponding to a 30-minute timestamp). The optimiser is required to select setpoints in the given range, including 0, to minimise the total cost of the day under the constraints. It can be predicted that running the cost function for every tested combination of setpoints in the model will inevitably lead to unnecessarily long run times. Therefore, the optimiser will cast 4 pre-calculated cost vectors for the whole dataset as inputs, the selection of which correspond to the nature of the constraints discussed. These input vectors are summarised below and can be calculated through simple manipulation of the cost function, where S_p denotes the setpoint of the engine in percentage units. Vectors of optimised costs for $50 < S_p \leq 100$ $(Costs_{50-100})$, $40 \leq S_p \leq 50$ $(Costs_{40-50})$, $S_p = 0$ $(Costs_{off})$ and $S_p = 100$ $(Costs_{100})$ are cast. Under correctly formulated constraints, and with the use of binary

variables, these vectors can be combined to output a final objective cost vector to be minimised.

ON/OFF Constraint

A vector of binary decision variables, On_off_i, is cast, with i ranging from 0 to 47. A value of 1 and 0 corresponds to the engine being on and off, respectively. *Eq. 8* represents a constraint allowing the reverse of this vector to be cast.

$$On_off_rev_i = -1 \times On_off_i + 1 \quad (8)$$

Using these binary variables with *Eq.* 9, a combination of $Costs_{50-100}$ and $Costs_{off}$ vectors can be formulated, giving an initial vector of optimised costs, Opt_costs_1.

$$Opt_costs_1 = Costs_{50-100} \times On_off_i + Costs_{off} \times On_off_rev_i \quad (9)$$

To model a switch between the engine being on or off, a vector is cast under *Eq. 10*,

$$On_off_switch_i = |On_off_i - On_off_{i+1}|, \quad (10)$$

where a value of 1 corresponds to an engine ON/OFF switch, and 0 otherwise. *Eq. 11* then models the Centrica recommendation of a maximum of 6 daily ON/OFF switches.

$$\sum_{i=0}^{46} On_off_switch_i \leq 6 \quad (11)$$

40-50% Constraint

Through the same strategy demonstrated by *Eq.* 9, casting a new set of 48 binary variables allows us to combine Opt_costs_1 with $Costs_{40-50}$ to create a second cost vector, Opt_costs_2, incorporating engine setpoints of 40-50%.

$$Opt_costs_2 = Opt_costs_1 \times On_off_forty_fifty_rev_i + Costs_{40-50} \times On_off_forty_fifty_i \quad (12)$$

$On_off_forty_fifty_i$ is 1 when the timestamp is chosen to operate with costs corresponding to a setpoint in the range of 40-50%, and 0 otherwise. $On_off_forty_fifty_rev_i$ is calculated in the same way as shown in *Eq. 8*. *Eq. 13 & 14* represent the constraints necessary for the model to function.

$$\sum_{i=0}^{47} On_off_forty_fifty_i \leq 6 \quad (13)$$

$$On_off_forty_fifty_i \leq On_off_i \quad (14)$$

Eq. 13 ensures that engine operation at 50% or below doesn't lie above 3 hours per day and *Eq. 14* ensures that time intervals for 40-50% operation can only be selected if the engine is on.

100% Output Constraint

The engine must operate at full output for at least 2 hours if 2 or more timestamps (30-minute intervals) are operated at 40-50% in one day. A vector of auxiliary variables is cast, $Full_output_aux_vector = [0, 0, 1, 1, 1, 1, 1]$, each value having a vector index equal to $\sum_{i=0}^{47} On_off_forty_fifty_i$ such that $Full_output_aux = Full_output_aux_vector\left[\sum_{i=0}^{47} On_off_forty_fifty_i\right]$. A third vector of 48 binary variables, $On_off_full_output_i$, is 1 when the corresponding value in the $Costs_{100}$ is to be implemented into the final cost vector, and 0 otherwise. *Eq. 15* makes the necessary constraint active when 2 or more timestamps have a setpoint less than 50%, and if so, ensures that 4 timestamps (2 hours) have costs corresponding to full output.

$$4 \times Full_{output_{aux}} \leq Full_{output_{aux}} \sum_{i=0}^{47} On_off_full_output_i \quad (15)$$

The final optimised cost vector can then be cast through *Eq. 16*,

$$Opt_costs_{final} = Opt_costs_2 \times On_off_full_output_rev_i + Costs_{100} \times On_off_full_output_i, \quad (16)$$

where $On_off_full_output_rev_i$ is formulated using the same strategy as in *Eq. 8*. $\sum Opt_costs_{final}$ is then the objective function to be minimised in the model. Input vectors corresponding to each iterative day in the relevant dataset are injected into the optimiser with the daily cost being minimised; total cost for the data set being calculatable by appending the objective for each day to a separate total cost vector.

Additional Details of the Pyomo Model

The solver used is 'mpec_nlp', one of the included free solvers in the Pyomo 'SolverFactory'. Run time for each daily dataset is approximately 5 seconds; the run time for optimisation of a year's data is roughly 30 minutes. Run time will therefore be about 3 hours when running the model for 6 different carbon taxes. Obtaining results for all 15 stores for the 20 carbon tax values investigated is associated with a total run time of 150 hours (Intel Core i7-6600U CPU, 8GB RAM). The '100% Output Constraint' however, is not included in the final optimiser applied to the stores due to an unnecessary increase in complexity when this constraint is active. For all days in the

Wellesbourne Sainsbury's dataset, there is always at least 2 hours at the end of the day where the engine ramps up to full output anyway. Other stores tested, which have much higher power demands are also seen to operate at full output for at least 2 hours each day. This means the addition of the 48 $On_off_full_output_i$ binary variables adds an unneeded additional level of complexity to the model, ramping the run time up to 2 minutes for a single day's data.

Results and Discussion

Key Savings results

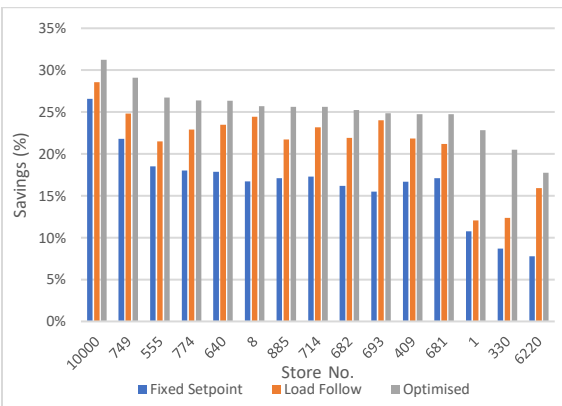


Figure 4: Percentage savings of Sainsbury's facilities relative to Annual Offline Cost

Fig. 4 shows the savings of the store portfolio for optimised, fixed setpoint and load following strategies. Savings are calculated as a percentage relative to the Annual Offline Cost (AOC), the energy cost without use of a CHP engine, importing electricity from the grid and supplying heat from a conventional gas boiler. Gross annual savings are found in *Table A in Appendix B*. Cost savings accrued vary significantly between stores, however for every store, the optimiser outperforms the load following strategy, which outperforms the fixed setpoint strategy. As established through preliminary analytics, adding operational flexibility improves the engine performance. Allowing the load following strategy to react to the demand profile of a store provides a great cost saving advantage over the fixed setpoint strategy. The additional flexibility of the optimiser provides an even greater advantage, increasing savings further.

Store Profiling

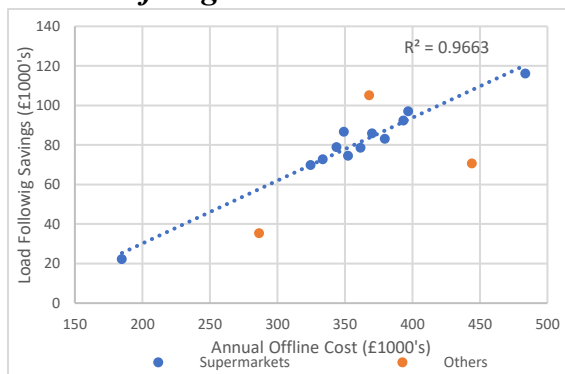


Figure 5: Load following strategy savings against annual offline cost

A correlation was found between load following savings of supermarkets and AOC. AOC was found to be a good metric to describe store 'size' as its essentially a weighted sum of heat and electricity demands. A simple shortcut can be applied to estimate this cost using *Eq. 17*, derived utilising regression and curve fitting techniques. The average annual electricity and heat demands ($\overline{x_E}$ and $\overline{x_D}$) in kWh can be used to calculate AOC within 2% of the value produced from the cost function.

$$AOC\ (£) = 414(\overline{x_E} + 4.2\overline{x_D}) \tag{17}$$

Fig. 5 shows that as store size increases, savings from the load following strategy increase linearly. Using the linear trend equation, a theoretical maximum percentage saving can be calculated. This was found to be 24% at £420,000 annual offline cost, a gross saving of over £100,000 per year. The three 'non-supermarket' facilities do not follow this same trend. The data points for the two distribution centres fall below the trendline, therefore accruing fewer savings than a supermarket of the same size. The data point for the corporate office falls above the trendline, therefore shows the opposite trend.

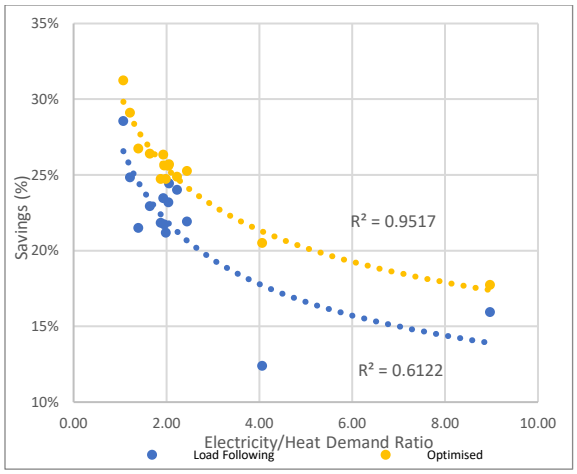


Figure 6: Percentage savings relative to AOC against electricity/heat demand ratio

Fig. 6 shows that percentage savings increase exponentially as the electricity/heat demand ratio decreases. This correlation can be explained using the CHP engine characteristics. The average engine electricity/heat production ratio is approximately 0.75. The closer a demand ratio is to this number, the greater the energy use efficiency will be. From our data set, the average supermarket operates with a demand ratio of approximately 2. This means that if the engine is operating to meet the electricity demand, only 37.5% of the heat produced is being utilised.

The outliers identified in *Fig. 5* are extreme data points in *Fig. 6***.** The points for the distribution centres lie on the far-right side of the graph with demand ratios of approximately 4 and 9, while the corporate office lies to the far left with the lowest demand ratio of approximately 1. This data explains why these facilities do not follow the savings trend identified in *Fig. 5*. With these extreme demand ratios, they utilise much more or much less of the heat produced from the engine than your typical supermarket, giving an explanation of the different levels of savings achieved.

The graph demonstrates that the difference in savings for load following and optimised strategies vary significantly between stores. Using the two distribution centres examples, the savings difference between strategies is over 4 times greater for one facility than the other. *Fig. 7* was produced to investigate this observation. The graph shows that as store size increases, the additional savings from the optimiser described in *Eq. 18* decreases. Using the fuel consumption, calculated from *Eq.* 1, we can deduce that the engine is more efficient at higher outputs, and therefore, at high setpoints the offline cost is usually much greater. It should be considered that when demand drops below the minimum engine output, switching the engine off can often become more favourable. The engine cannot run below minimum output, so more electricity must be produced than demanded. This surplus is usually exported at a price lower than the engine production cost. For these two reasons, larger stores with higher demand benefit much less from the ability to turn off. Furthermore, during the few occasions where export prices are high, if the load follow setpoint is already near maximum output, the additional benefit from ramping up production is reduced.

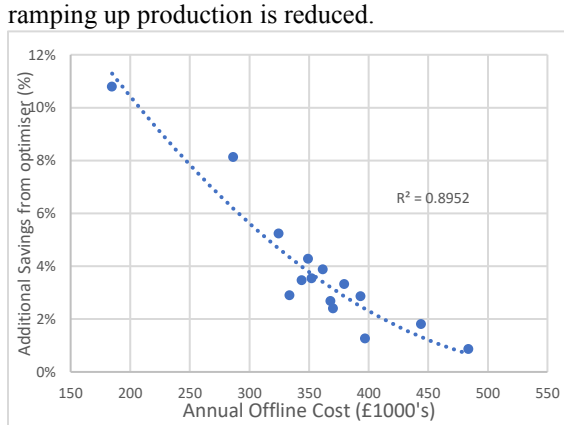


Figure 7: Additional savings from the optimiser against annual offline cost

$$Savings_{Additional} = Savings_{Optimised} - Savings_{LF} \quad (18)$$

Environmental Analysis

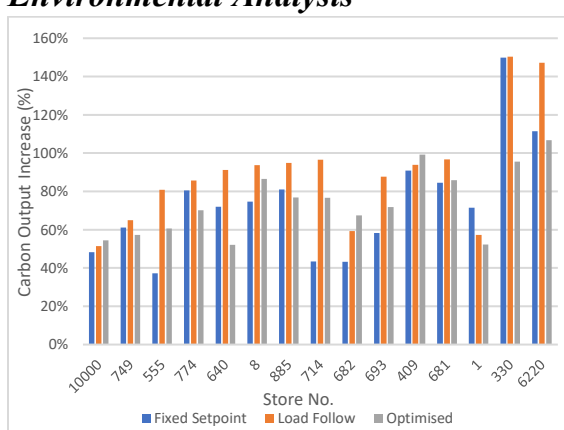


Figure 8: Percentage carbon output increase of Sainsbury's facilities relative to operation without CHP engine

Unfortunately, with modern UK grid electricity having such a low carbon intensity (National Grid

ESO, 2019), the nature of the engine means irrespective of the strategy followed, increases the stores carbon footprint are seen, demonstrated in *Fig. 8*. The environmental impact is highly variable between stores. The argument could be made that running the engine will reduce electricity demand from the marginal producers on the grid, namely CCGT (combined cycle gas turbines). Comparing engine emissions with the grid average carbon intensity, shown in *Fig. 8,* could therefore be considered an unfair assessment. When compared with CCGT emissions, (Mearns, E, 2016) the average carbon emissions for all strategies is lower due to the engines higher fuel efficiency.

The additional carbon output is particularly high for the two distribution centres (store ID's 330 and 6220). As expected, the previously explained energy inefficiencies are detrimental to environmental performance. The more energy wasted, the more carbon dioxide is released per unit of energy utilised. The relationship of carbon output between operational strategies is much less conspicuous than that of cost savings. However, on average, the fixed setpoint strategy has a best environmental performance, closely followed by the designed optimiser, and finally the load following strategy. The load following emissions are higher than that of the fixed setpoint strategy for all but one store. The chosen fixed setpoint of 162kWe is lower than the average demand of most of the stores. The total power generated will be greater for the load following strategy and therefore more fuel will burned and more carbon dioxide released. The only two facilities with electricity demand ratios below 162kWe are store ID's 1 (Wellesbourne) and 330. The Wellesbourne store has a much lower demand than any other facility tested, explaining why fixed set point emissions are higher than for load following. Store 330 deviates from this, showing fixed setpoint emissions that are slightly lower. This is due to a large amount of electricity being generated at low engine output and thus, a much lower efficiency. This means the amount of fuel burned overall will be higher for the load following strategy and more carbon dioxide will be released.

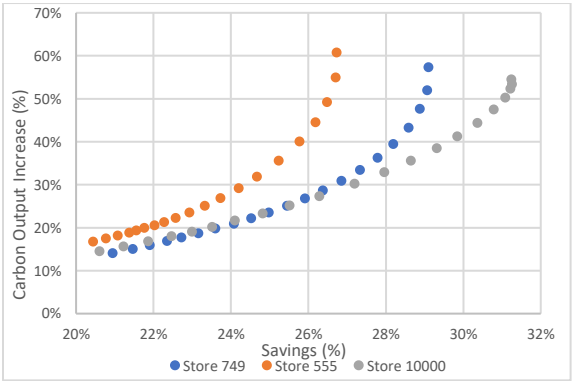


Figure 10: Percentage carbon output increase against percentage savings of the optimiser from 0 to 200 £/tCO2 carbon tax

Carbon Tax (£/tCO2): 0

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
00:00:00	197	198	197	196	199	193	196
01:00:00	250	197	196	197	200	193	195
02:00:00	250	194	196	198	205	199	197
03:00:00	196	194	198	205	200	192	153
04:00:00	143	189	194	192	189	190	155
05:00:00	143	193	188	193	181	189	141
06:00:00	146	186	188	191	157	184	162
07:00:00	147	196	180	185	0	185	158
08:00:00	156	198	176	179	0	179	164
09:00:00	183	198	179	180	0	177	180
10:00:00	170	202	173	172	125	172	174
11:00:00	165	187	161	165	125	163	161
12:00:00	158	159	157	125	125	137	159
13:00:00	126	152	143	125	125	125	135
14:00:00	148	167	150	125	125	125	127
15:00:00	149	173	159	125	125	125	126
16:00:00	153	158	152	125	125	125	126
17:00:00	250	250	250	250	250	250	180
18:00:00	250	250	250	250	250	250	150
19:00:00	250	250	250	250	250	250	163
20:00:00	199	196	194	130	188	191	192
21:00:00	207	211	196	134	202	213	200
22:00:00	200	199	197	151	197	193	194
23:00:00	191	195	197	160	197	193	193

Carbon Tax (£/tCO2): 90

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
00:00:00	147	199	183	165	199	173	176
01:00:00	147	197	173	196	200	193	150
02:00:00	150	194	178	147	191	198	144
03:00:00	142	196	169	146	178	167	153
04:00:00	143	195	168	137	184	163	155
05:00:00	0	196	159	138	175	147	141
06:00:00	0	196	163	148	136	185	161
07:00:00	146	196	174	152	0	178	157
08:00:00	155	199	175	178	0	174	163
09:00:00	164	199	179	180	0	177	164
10:00:00	143	203	173	171	125	173	165
11:00:00	131	187	155	165	125	162	161
12:00:00	135	160	136	0	125	0	135
13:00:00	143	152	144	0	125	0	136
14:00:00	148	167	150	0	0	125	142
15:00:00	150	174	159	125	0	0	141
16:00:00	154	158	152	0	0	0	143
17:00:00	250	250	250	250	250	250	147
18:00:00	250	250	250	250	250	250	155
19:00:00	250	250	250	250	250	250	164
20:00:00	199	196	194	130	188	191	192
21:00:00	207	211	196	134	202	213	200
22:00:00	200	199	196	151	196	193	194
23:00:00	191	195	196	160	196	193	193

Carbon Tax (£/tCO2): 180

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
00:00:00	147	199	183	165	199	173	176
01:00:00	146	197	173	151	196	174	150
02:00:00	150	194	177	147	190	164	144
03:00:00	0	196	169	145	177	166	153
04:00:00	0	195	168	137	184	163	0
05:00:00	0	196	159	138	175	147	0
06:00:00	0	196	162	148	136	186	161
07:00:00	146	197	173	151	0	178	157
08:00:00	155	199	169	138	0	174	163
09:00:00	164	199	168	125	0	169	163
10:00:00	143	250	250	128	0	173	165
11:00:00	130	188	155	125	0	137	131
12:00:00	134	250	135	0	0	0	135
13:00:00	143	250	144	0	0	0	136
14:00:00	148	167	151	0	0	0	142
15:00:00	150	174	159	0	0	0	141
16:00:00	154	158	152	0	0	0	144
17:00:00	250	250	250	250	250	250	147
18:00:00	250	250	250	250	250	250	155
19:00:00	250	250	250	250	250	250	164
20:00:00	157	250	160	125	250	250	178
21:00:00	156	197	174	0	250	176	179
22:00:00	176	250	176	0	166	170	166
23:00:00	190	190	168	250	159	159	178

Figure 9: Heat maps showing optimiser setpoints for Store 749 at different levels of carbon tax

In order to reduce emissions from the optimised strategy, multi-objective optimisation can be utilised by introducing an artificial 'carbon tax'. *Fig. 9* demonstrates that a large improvement in environmental performance can be achieved by sacrificing a small amount of savings. Across the store portfolio, the steeper sections of the curves in *Fig. 9* shows additional carbon output can be halved with an average savings reduction of just 10%.

To explain how the optimiser achieves this result, heat maps are constructed showing the engine setpoints selected throughout the day for different values of carbon tax. From *Fig. 10*, there are 3 key observations. Firstly, as carbon tax increases from left to right, the amount of time the engine spends offline (shown in blue) increases. Cost savings are sacrificed by turning the engine off and importing low carbon electricity instead. Secondly, a less obvious trend is that higher setpoints are replaced with slightly lower setpoints, most easily seen by comparing the 0 and 90 £/tCO2 maps during the early hours of Thursday morning. Instead of meeting the store electricity demand, the optimiser selects a setpoint that meets the lower heat demand and the additional electricity required is imported from the grid. This ensures 100% of the energy produced by the engine is utilised, a much lower carbon solution. Finally, the high carbon tax heatmap shows timestamps outside of the peak export price hours where the engine is at maximum output. During these times, the carbon intensity of the grid is particularly high, surpassing the value that the CHP engine can achieve. The engine is then ramped up to 250kWe to export lower carbon electricity to the grid.

Sensitivity Analysis

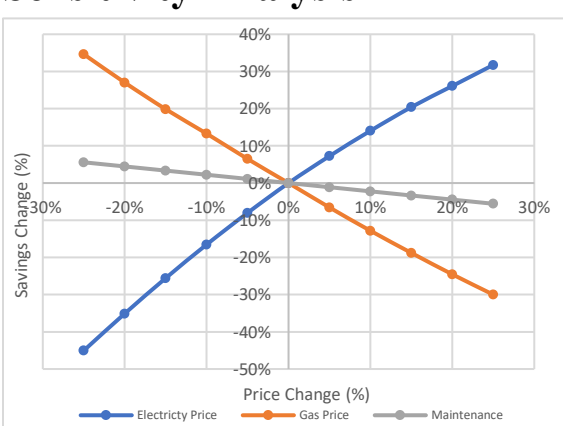


Figure 11: Percentage change of savings against percentage change of price variable

The results in *Fig. 11* demonstrate CHP engine savings are particularly sensitive to gas and electricity prices. A 20% increase in electricity price will increase savings by 26%, whereas a 20% increase in gas price will decrease savings by 25%. The results show that the effects of the two price variables are approximately equal and opposite. The same proportional change in both prices would result in a negligible change in predicted savings.

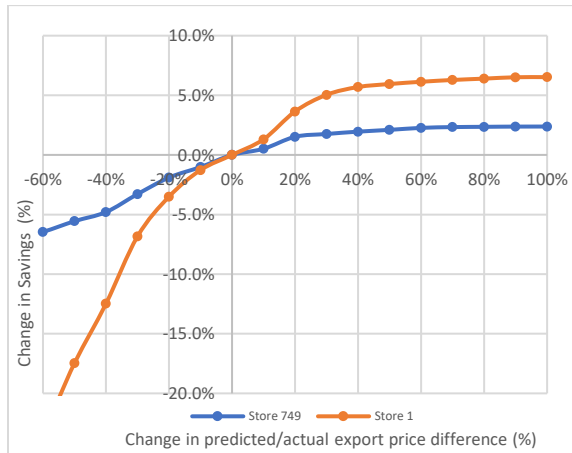


Figure 12: percentage change in predicted/actual electricity export price difference against change in savings

$$Price_{Diff} = Price_{actual} - Price_{Predicted} \quad (19)$$

$$Price_{Sens} = Price_{actual} - Price_{Diff}(1 - \frac{\theta}{100}) \quad (20)$$

Analysis was conducted to investigate how the percentage accuracy (θ) of the electricity export price predictions affected savings. *Fig. 12* shows that savings increase as the predicted price change moves towards 100%, where predicted price becomes equal to actual price. However, most of the savings potential is achieved after a 30/40% increase in accuracy. *Fig. 12* also demonstrates that small decreases in accuracy results in significant decreases in savings. More accurate data would allow the optimiser to make fewer mistakes and unlock the maximum potential savings as shown in *Fig. 12*.

As expected, changing price prediction accuracy has a much greater effect on savings for store 1 than the much larger store, 749. For a larger store, the load following setpoint will be closer to the engines maximum output; the surplus that can be produced will be smaller. This means larger stores will benefit less from the ability to ramp up to maximum output, thereby less affected by changes in price prediction accuracy.

Conclusions

Preliminary data analytics revealed significant cost benefits from widening the operational domain of the CHP engine. Flexible operation constrained by Centrica recommendations was built into an optimisation model, which was applied to historical data of multiple Sainsbury's facilities. The designed optimiser was shown to accrue greater savings than the two baseline strategies. This result increases the financial feasibility of engine installation in any of Sainsbury's facilities.

Store profiling revealed that supermarket savings from the load following strategy were shown to increase linearly against store size, with a maximum percentage savings of 24% identified at the AOC of £420,000. From this, we can conclude that supermarkets of this size are most suitable for executing a load following strategy with the 250kWe CHP engine. It was shown that for future analysis, stores of this approximate size can be identified using shortcut *Eq. 17*. From this, a preliminary evaluation of suitability can be made.

Percentage savings of the optimised and load following strategies were shown to increase exponentially as the electricity/heat demand ratio of a facilities increase. It was concluded that facilities with demand ratios closest to the average (0.75) will achieve the highest energy efficiency and benefit most from the CHP engine. It was also shown that the additional capabilities of the optimiser, relative to the load following strategy, were more beneficial for smaller facilities. If Sainsbury's were to install the 250kWe engine in a smaller store, utilising the benefits of the optimiser is more relevant for cost savings, and thus, economic feasibility.

Environmental analysis demonstrates that exercising multi-objective optimisation can significantly reduce carbon emissions from the engine with a small sacrifice in savings. Before engine installation, the acceptable cost/carbon trade off would need to be established and the corresponding value of carbon tax implemented.

Outlook

To improve the confidence of conclusions made, the data set analysed must be increased, testing a greater range of facilities. As explained in the report, the average electricity/heat demand ratio of the supermarkets tested is far from the ideal value. Identifying a facility type with a demand ratio closer to the engine production ratio would be extremely valuable.

Sensitivity analysis demonstrates that development of a more accurate export price prediction algorithm can help unlock further savings, the results of which could be applied in many settings outside the context of this project, making this work particularly valuable.

A comprehensive cost benefit analysis could be conducted for the engine and optimiser installation. Sensitivity analysis showed that savings are particularly sensitive to gas and electricity prices, an in-depth cost projection would need to be developed to assess the true economic feasibility of installation.

A potential area for investigation is the procurement and use of biogas. Contracts can be purchased to inject sustainable biogas into the network, offsetting the carbon emissions of any boiler or engine gas combustion (GGCS, 2019). A fuel switching capability could be built into the model providing an alternative method for reducing carbon emissions.

References

Yining, L. (2018) 'Combined Heat and Power (CHP) Control Strategies Under Uncertainty', *Energy Futures Lab,* Imperial College London

Mago, P. J. *et al* (2010) 'Analysis and optimization of the use of CHP-ORC systems for small commercial buildings', *Energy and Buildings*, Mississippi State University

Alvarado, D. C. *et al* (2016) 'A Technology Selection and Operation (TSO) optimisation model for distributed energy systems: Mathematical formulation and case study', *Applied Energy*, Department of Chemical Engineering, Imperial College London

Exelon (2019). Available at: http://www.elexon.co.uk

National Grid ESO (2019). Available at: http://carbonintensity.org.uk

Green Gas Certification Scheme (2019), Renewable Energy Assurance Limited. Available at: http://www.greengas.org.uk/scheme

Mearns, E. (2016) 'CO2 Emissions Variations in CCGTs Used to Balance Wind in Ireland', *Energy Matters*. Available at: http://euanmearns.com/co2-emissions-variations-in-ccgts-used-to-balance-wind-in-ireland/

Carbon composite anodes in carbonate- and ether-based electrolytes for potassium-ion batteries

Shengming Zhang, Anastasia A. Teck
Department of Chemical Engineering, Imperial College London, U.K.

Abstract Potassium ion batteries appear promising in addressing the need for sustainable and cost-effective large-scale energy storage, particularly with fluctuating renewable energy supply and demand in mind. A major challenge in this regard is to understand and improve anode and electrolyte material compatibility. In this work, pure and composite carbon anodes, in particular hard carbon – soft carbon and, for the first time, graphite – soft carbon composites, were explored in carbonate- and ether-based non-aqueous electrolytes. A systematic investigation of a range of composite ratios is presented, with a subsequent detailed study of three selected material combinations. It is found that graphite – soft carbon composites show high initial coulombic efficiencies up to 67.3% and broad working potentials around 0.5V vs K^+/K in carbonate-based electrolytes, as the presence of soft carbon inhibits solid electrolyte interface layer formation and provides a flexible protection to the graphitic layers. Compared to graphite, the hard carbon – soft carbon composite shows improved rate performance in the ether-based electrolyte, but surprisingly poor cyclability. Importantly it is discovered that graphite in the ether-based electrolytes exhibits instability with both intercalation and solvent-co-intercalation potassium storage.
Keyword: Potassium-ion battery, graphite-soft carbon, hard carbon-soft carbon, carbonate- and ether-based electrolyte.

1. Introduction

Batteries are required for a range of applications in modern day life, from portable electronic devices and vehicles to stationary electrical energy storage for the integration of intermittent renewable energies, such as wind and solar, into the grid[1]. Their development is critical to reduce dependence on fossil fuel and address the climate crisis. For this, cost effective, efficient, and large-scale energy storage is required[2].

Room temperature, rechargeable lithium ion batteries (LIBs) are widely used for portable energy storage, due to a high charge-to-weight ratio allowing for a high energy density and long cycling stability. Many commercial LIBs use graphite anodes, with a theoretical capacity of 372mAh/g, and $LiCoO_2$ cathodes[3]. However, sustainability is an issue for LIBs. Lithium has an uneven global distribution and a low abundance of just 0.0017wt% in the earth's crust, resulting in rising prices and the risk of exhausting reserves[4]. With other critical materials, such as Cobalt, also experiencing rising prices[1] and LIBs having using expensive Copper current-collectors instead of Aluminium to avoid electrochemical side reactions, researchers are exploring alternative metal-ion batteries to ensure low acquisition costs and continued availability[5].

Elements of the same periodic group display similar physico-chemical properties, indicating that other Alkali metals may be suitable replacements for lithium. The next two alkali metals are Sodium and Potassium, which are relatively abundant, with 2.3wt% and 1.5wt%, respectively[6], in the crust. Sodium-ion batteries (NIBs) were expected to provide a similar performance to LIBs and have therefore received wide-spread attention. However, while research into NIB cathode materials has been quite successful, the anode side has been problematic[3], as sodium ions do not electrochemically intercalate well into graphite at room temperature and require expanded graphite or solvent co-intercalation to achieve stable graphite intercalation compounds (GICs) of high sodium content[2]. Therefore, as research into suitable NIB anode and electrolyte materials is on-going, some researchers have recently started looking at potassium-ion batteries (KIBs). Pramudita[3] explain that in its infancy, KIB research appears promising with a comparable performance to LIBs and a similar graphite intercalation mechanism, achieving reversible capacities greater than 200mAh/g. KIBs are especially interesting for large scale energy storage due to the high ionic conductivity and small Stokes radius of K^+ [7] promising a high power density, as well as cheaper electrolyte materials.

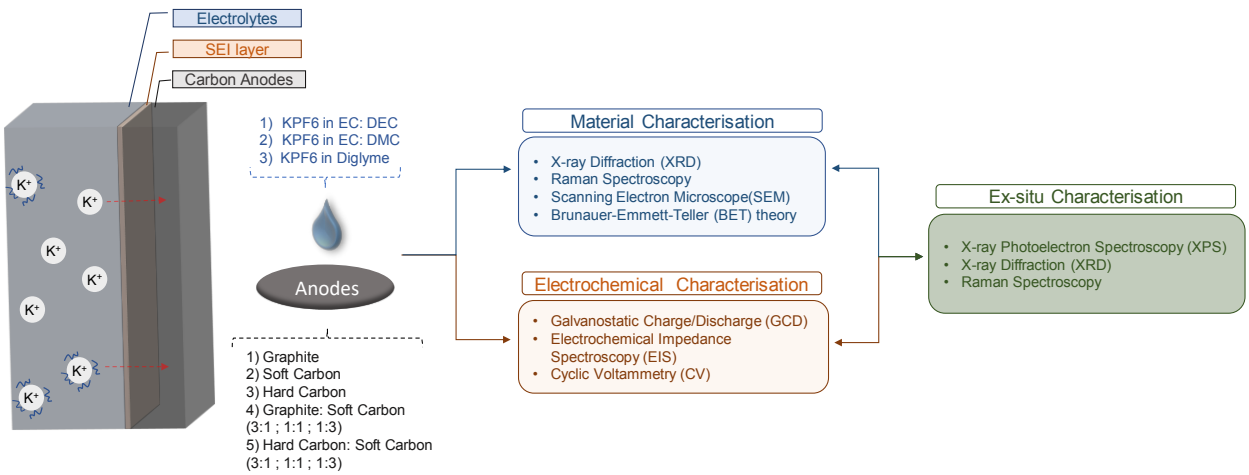


Fig. 1 – Schematic Abstract

The performance of KIBs is highly dependent on the choice of materials, such as for the electrodes and electrolyte, which consists of a salt, typically KPF_6, and an aqueous or non-aqueous solvent. A range of successful cathode materials have been reported, including Prussian blue analogues, layered oxides and polymers[8-14]. To further improve KIB performance, researchers have investigated anode materials beyond graphite, including non-graphitic carbons, such as Soft Carbon (SC)[15] and Hard Carbon (HC)[15-17], metals and metal-oxides[18-20].

Graphite is the dominant anode in LIBs due to its high practical specific capacity of 360mAh/g (theoretical 372mAh/g) and low operating potential of 1V vs Li^+/Li and has been widely investigated for KIBs as well[2]. The structure of graphite is well understood, and researchers generally agree that the electrochemical intercalation of potassium follows the stages → KC_{48} (stage 4) → KC_{36} (stage 3) → KC_{24} (stage 2) → KC_8 (stage 1) during potassiation[2, 21]. Graphite has previously achieved a specific capacity of 273mAh/g (theoretical 279mAh/g)[6] and is convenient due to its low cost, abundance and environmentally friendly nature[21]. However, it is thought that potassiation causes large volume expansion of over 60% due to the size of K^+, such that graphite shows poor cycling performance and fast capacity fading[2].

Soft carbon is a partially disordered, semi-graphitic carbon, with relatively few defects and high crystallinity[22], that forms graphite if heated. Fan, Ma[21] describe the structure as divided into a high strain region with disorganised carbon and a low strain, graphitic region, resulting in a high electrical conductivity[22] and a significantly lower density compared to graphite, such that soft carbon exhibits a very high rate capability. Due to the disordered structure, various storage mechanisms, including intercalation, adsorption, pore filling, and defect interactions, should in theory allow the capacity to exceed that of graphite[23].

Hard carbon is a non-graphitisable carbon, consisting of graphene-like carbon layers and micropores between disordered carbon stacks[24]. Like soft carbon, it has a larger interlayer spacing than graphite, allowing for better rate performance, and a disordered structure allowing for a combination of multiple storage mechanisms[25]. Most notably, hard carbon was found to have a high cycling ability, which is attributed to its cross-linked structure[25] and consequently unchanged structure after potassiation[16].

To combine the benefits of both hard and soft carbon, Jian[15] produced a HC-SC composite with 20wt% SC as a KIB anode, which achieved significantly better rate performance than pure HC, while maintaining its stable cyclability. The smaller SC particles were spread over the HC matrix, plugging its surface pores, thereby decreasing the irreversibility of potassiation[15]. Xie, Xu[26] further tested HC-SC composites for sodium storage and state that investigating composites of different ratios would allow tuning of the structure to improve guest particle storage in the host material. While some HC-SC composites have been investigated for KIBs, and the possibility of graphite-HC (G-HC) composites has been explored for LIBs[24], no literature was found on graphite-SC (G-SC) composites, which promise higher working potentials and better rate performance than graphite and a lower cost than composites with HC.

The compatibility of electrode and electrolyte is essential in the development of KIBs, as the electrolyte can impact the potassium storage mechanism[27], as well as the extent of formation and stability of the solid electrolyte interface (SEI) layer[2]. A range of non-aqueous electrolyte solvents have previously produced high reversible capacities in KIB anode materials[6]. A mixture of ethylene carbonate and diethyl carbonate (EC-DEC) has been used for a variety of anode materials, including SC and HC, with promising cycling performance[3]. Recently, Wang[27] performed a study comparing carbonate- and ether-based electrolytes, which found that the dimethoxyethane (DME) based electrolyte resulted in lower volume-expansion, higher diffusion rates, and higher insertion voltages than the di-methyl carbonate (DMC) based electrolyte, suggesting that co-intercalation lowered the destructive interaction force with the graphite anode. Consequently, different electrode materials are expected to perform best with different solvents.

In this work, we present a systematic investigation of different ratios of HC-SC and, for the first time, G-SC composites (25wt%, 50wt%, 75wt%) in three different non-aqueous carbonate- and ether-based electrolytes (KPF_6 in EC-DMC, EC-DEC, Diglyme), as visualised in **Fig. 1**. Using material characterisation of pristine samples and galvanostatic charge-discharge (GCD) tests of half cells we perform an overview of all materials and select three for further electrochemical and ex-situ testing after cycling.

2. Methods

2.1 Material synthesis:

Hard Carbon (HC) was obtained by annealing Hydrothermal carbonisation (HTC) powder at 1300 °C for 2 hours under N_2 flow with an initial heating rate at 5 °C min^{-1} until reaching the target temperature. The HTC powder was synthesised from D-glucose powder (5g) treatment in deionised water (45g) in a closed Teflon reactor at 200°C under self-generated pressure for 12 hours. The produced black precipitate was filtered and rinsed with DI water until the pH of washing effluent was 7, then the washed cake was dried in a vacuum oven at 80 °C overnight. Graphite Powder was purchased from Sigma-Aldrich and used without further modification. Soft Carbon (SC) was prepared via pyrolysis of mesophase pitch (Bonding Chemical, USA) in a tube furnace at 1100 °C for 2 hr under N_2 atmsophere. All different compositions were synthesed by mixing graphite, HTC powder and uncarbonised pitch in the desired mass fractions before the annealing process. The mixtures were continuously ground for 30 minutes in the agate mortar and pestle set to obtain blended fine powders, which were annealed in a tube furnace at 1100 °C for 2 hr under N_2 flow with an initial heating rate of 5 °C min^{-1} until the target temperature was reached.

2.2 Electrochemical Characterisation:

The working anodes were prepared by mixing and milling 80 wt% of selected sample composition, 10 wt% of carbon black (Alfa-Aesar) to improve conductivity and 10 wt% of polyvinylidene difluoride (PVDF, from Sigma-Aldrich) as the binder in N-Methyl-2-pyrrolidone (NMP, from Sigma-Aldrich). The homogenous slurry was pasted onto Al foil (MTI Corporation) for first-stage mild drying at 50 °C for 2 hr, ensuring the even deposition of a dense active material layer. The dried anodes were cut into square pieces (Area of 0.64 cm^{-2}) with approximately 2 mg cm^{-2} active material loading, and the mass of individual anode was measured and recorded. The pre-dried anodes were further desiccated at 105 °C under vacuum overnight and transferred immediately into an Argon-filled glovebox (mBraun, H_2O, O_2 < 0.5 ppm). In the half cell configuration, Potassium metal foil with a diameter of 10 mm was used as the counter and reference electrode at the same time, and the glass fiber paper (Whatman, GF/D Glass microfiber filters) with a diameter of 16 mm served as the separator. The three electrolytes used were 0.8 mol L^{-1} KPF_6 in ethylene carbonate and diethyl carbonate (EC: DEC 1:1v%), 0.8 mol L^{-1} KPF_6 in ethylene carbonate and dimethyl carbonate (EC: DMC 1:1v%) and 0.8 mol L^{-1} KPF_6 in diglyme. The stainless-steel coin cells (MTI Corporation, CR2032) were compacted using a gas driven crimper (MTI Corporation) at a gauge pressure of 120 Psi. All the electrochemical measurements were conducted at room temperature and pressure. Galvanostatic charge-discharge (GCD) profiles at interval of 0-2.5V (vs. K/K^+) and rate/cycling performance were recorded on a LAND CT 2001A battery testing system at different current densities. Cyclic voltammetry (CV) at 0.1 mV s^{-1} and electrochemical impedance spectroscopy (EIS) from 100 kHz to 0.01 Hz were measured on an Autolab PGSTAT204 electrochemical workstation.

2.3 Structural Characterization:

For powder samples, materials were dried at 105 °C in a vacuum oven overnight before performing structural characterisation. For ex-situ Characterisation, coin cells were unpacked using the crimper with the disassembling module after one complete cycle of discharging and charging with extra discharging at 0.1 C. The anode was removed from the stainless-steel wafer and washed with the corresponding electrolyte. Before ex-situ structural characterisation, washed anodes were dried at 105 °C in a vacuum oven overnight. X-ray photoelectron spectroscopy (XPS) was conducted by using a Thermo Fisher K-Alpha$^+$ XPS system, for which data was fitted using the Avantage Software. Raman Spectroscopy was performed on a Thermo Scientific DXR3 instrument and results were fitted using the Lorentzian method. The morphology of samples was obtained using a LEO Gemini 1252 FEG-SEM. X-ray Diffraction (XRD) was performed in an X'Pert PRO PANalytical using 40 mA and 40 kV power setting. The Brunauer-Emmett-Teller (BET) method and Density-functional-theory (DFT) method were used to evaluate surface area and pore size distribution from N_2 adsorption/desorption isotherm in Micromeritics 3 Flex Physisorption.

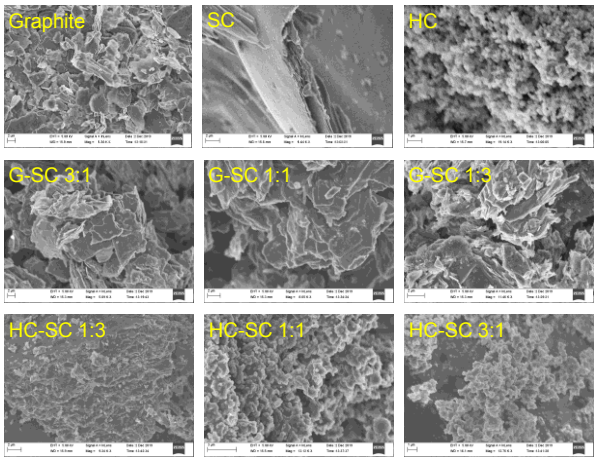


Fig. 2 – SEM of pristine anode active materials

3. Results

3.1 Material Characterisation

Scanning electron microscopy (SEM) discloses the morphology of pure graphite, soft carbon, hard carbons and their composites in **Fig. 2**. The pure graphite has a typical flake shape, which is covered by mesoporous pitch in G-SC composites after carbonisation, where the soft carbon serves as a glue to connect separate graphite flakes and creates continuous structure. The morphology for hard carbon is shown to be spherical particles with diameter ranging from 0.1m to 0.5 um. The role of soft carbon is similar in HC-SC composites to bind dispersive spherical partials together. The surface morphology and cross-sectional area for prepared anodes are demonstrated in **Fig. S6**. The carbon black is dispersed in the structure with binding of the PVDF network. The thickness of aluminium foil and the entire deposit of carbon mixture is around 10 µm and 50 µm respectively. The thickness and active material loading were kept at similar levels for all prepared anodes.

From Raman Spectroscopy, the defect and disorder structure of synthesised materials were determined by curve fitting and the I_D/I_G ratio. I_D/I_G for pure graphite powder, pure soft carbon powder and hard carbon powder is 0.12, 1.04 and 1.06 respectively. According to the three-stage model for graphite and amorphous carbon[28-30], pure graphite and soft carbon are classified in stage one, representing the transition from graphite to the nanocrystalline graphite. The I_D/I_G ratios for G-SC 9:1, G-SC 3:1, G-SC 1:1 and G-SC 1:3 are 0.3, 0.36, 0.53 and 0.68 respectively, indicating an increasing amorphousness as I_D/I_G ratio rises, which closely correlates to stage one behaviour. The hard carbon is typically allocated in stage two with broad D and G peaks[31]. All HC-SC composite have an I_D/I_G ratio of 1.03 and similar density of defects.

3.2 *Overview Study*

GCD tests were conducted at rates from 0.1C to 2C and back to 0.1C, where 1C is defined as the current rate at which it takes 1h for KC_8 to form in the carbon anode. The capacity, coulombic efficiency and rate capability data from these tests were used to gain a brief overview of the different electrode - electrolyte combinations, to identify interesting trends and draw preliminary

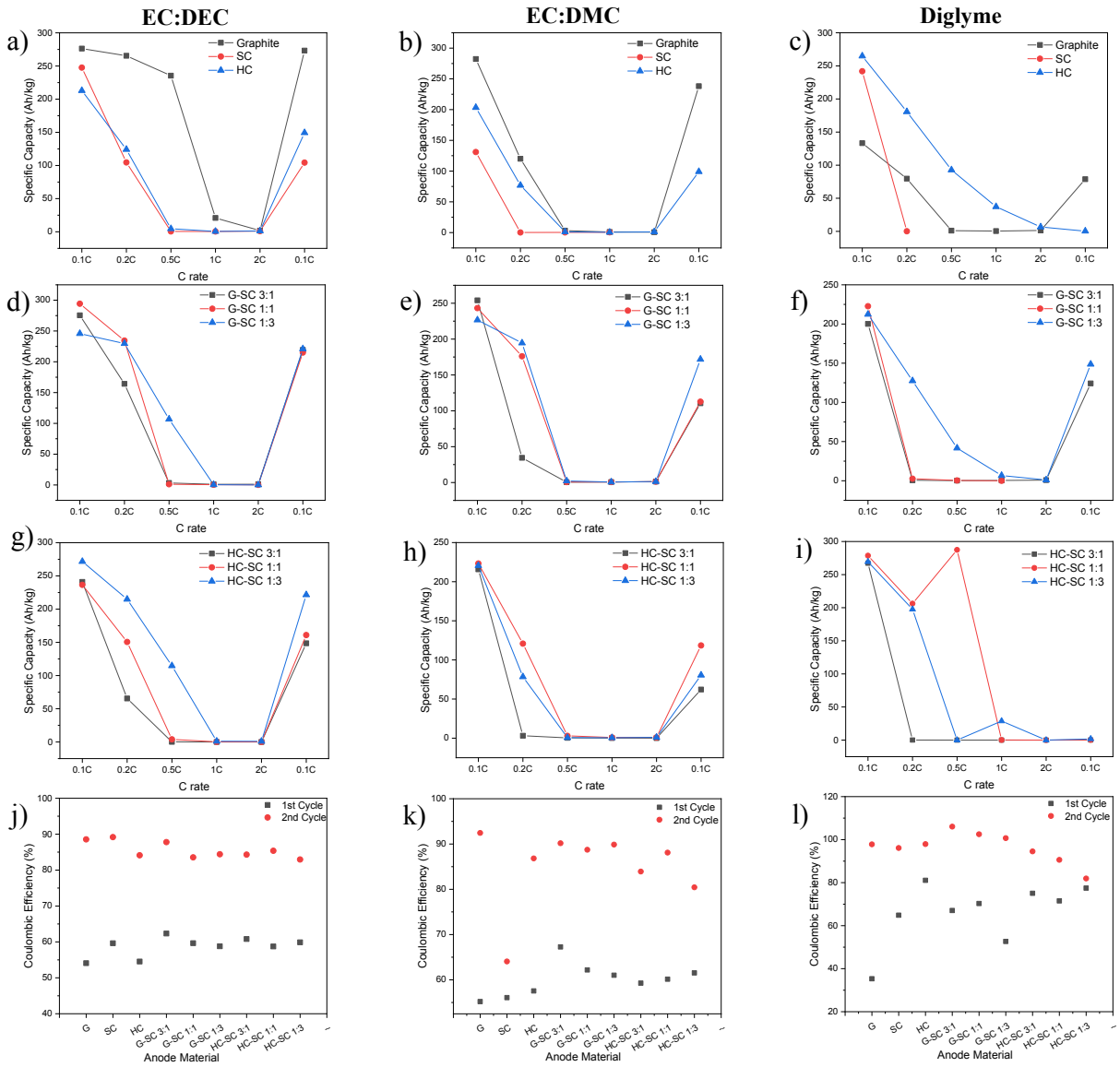


Fig. 3 – comparative rate performance based on GCD second cycle depotassiation capacities for graphite, SC, and HC in a) EC:DEC, b) EC:DMC, and c) Diglyme; for G-SC 3:1, 1:1, and 1:3 in d) EC:DEC, e) EC:DMC, and f) Diglyme; for HC-SC 3:1, 1:1, and 1:3 in g) EC:DEC, h) EC:DMC, and i) Diglyme; as well as 1st and 2nd GCD cycle coulombic efficiencies at 0.1C of 9 pure and composite anode materials in j) EC:DEC, k) EC:DMC, and l) Diglyme

conclusions. This initial understanding was then used to identify three samples to study further in detail in 3.3.

3.2.1 EC-DEC electrolyte

The nine anode materials are first compared in the carbonate-based electrolyte containing EC:DEC. Due to SEI layer formation in the 1st cycle, the more stable 2nd cycle data is used to contrast capacities, as shown in **Fig. 3** and **Fig.** S9. Among the pure materials (graphite, SC, and HC), graphite exhibits the highest second cycle depotassiation capacity of 276mAh/g, which is impressively close to the theoretical KC_8 capacity of 279mAh/g. However, both SC and HC samples have comparatively low capacities of 248mAh/g and 213mAh/g, respectively. While one would expect these more disordered carbons to exceed the capacity of graphite due to additional K storage mechanisms, Jian et al.[15] found a similar value of 246mAh/g for SC, but a higher value of 260mAh/g for HC. The G-SC composite materials underwent a maximum capacity of 294mAh/g at the 1:1 ratio, exceeding the capacity of both graphite and SC. This indicated that graphite and SC have a optimum increase of active sites at their interface when combined, depending on the ratio. Considering the HC-SC composites, the 1:3 ratio sample reach the highest capacity of 272mAh/g, while the other composites show only moderate capacities. As Xie et al.[26] found that HC-SC composites undergo an optimum ratio for maximum capacities in NIBs, due to a balance of SC blocking HC pores against an increase in ordered structures, it is possible that HC-SC reaches this optimum at high SC ratios, and would warrant further investigation of additional composite samples.

Fig. 3j) presents the 1st and 2nd cycle coulombic efficiencies. All samples in EC:DEC show exceptionally similar 1st and 2nd cycle efficiencies ranging between 54.1-62.3% and 83.0-89.2%, respectively. This similarity indicates that the extent of formation and stability of the SEI layer is uniform in all samples, such that SEI layer formation seems to be primarily governed by the electrolyte, and less the

characteristics of the anode material. The magnitudes of the efficiencies are similar to previously reported values for carbon anodes with PVDF binder[15]. As seen in Figure SM-X, the low initial coulombic efficiencies result from a large irreversible capacity and SEI layer formation in cycle 1. Interestingly, G-SC 3:1 exhibits the highest 1st cycle efficiency, with a relatively high 2nd cycle efficiency of 87.8% as well. This finding will be investigated further in 3.3.

The rate performance, as seen in **Fig. 3**, is relatively poor, with all samples eventually decreasing to zero capacity by a rate of 2C. All samples partially recover capacity when returned to 0.1 C, demonstrating that the rate performance is a result of the slow kinetics of K storage, rather than structural damage due to varying current densities. However, most samples do not recover completely, due to poor cyclability, as can already be seen in early cycles in **Fig. S10** and **Fig. S13**. In EC:DEC, graphite shows the best rate capabilities, only decreasing to 2mAh/g at 2C, and approximately maintaining a capacity to 237mAh/g at 0.5C. It also recovers very well to 99% of its original capacity, with **Fig. S13** showing that graphite is very stable over 30 cycles. The capacity of SC and HC decrease roughly linearly to zero by 0.5C, with SC recovering worse due to its extremely poor cyclability as shown in **Fig. S13**. This result is unexpected, as Jian[6] had found SC to have a superior rate capability to both graphite and HC, but does agree that HC is more stable than SC for cycling[25]. Among the G-SC composites, the 1:3 ratio gives the best rate capabilities, maintaining relatively constant capacities at all low current densities, but still decreases to zero at 1C. Interestingly, all of the G-SC composites recover to the same capacity. Coinciding with its high initial capacity, HC-SC 1:3 shows a reasonably good cycling performance, but exhibits a linear decrease, while the other HC-SC composites show the worst rate performance in EC:DEC, indicating that the addition of SC improves the rate capability notably, as expected.

3.2.2 EC-DMC electrolyte

GCD reveals that EC:DMC achieves lower 2nd cycle charge capacities than EC:DEC in all materials other than graphite. From **Fig. 3b)**, the graphite sample reaches 282mAh/g, which lies marginally above the theoretical capacity of KC_8 intercalation and may be attributed to inevitable uncertainty in experimental testing but could potentially indicate that a storage mechanism other than intercalation, such as adsorption on defects in the graphite structure, occurs. SC had the lowest capacity of 131mAh/g, suggesting that the choice of electrolyte has a great effect on this material, possibly through sluggish kinetics and irreversible potassiation. HC has a capacity of 203mAh/g, slightly below that in EC:DEC, appearing to be less significantly affected by the choice of electrolyte solvent. The capacity of the G-SC composites peaks at the highest graphite content (3:1) with 254mAh/g, unlike in EC:DEC, which implies that the optimum composition is shifted by the electrolyte, most likely in relation to solvent stability and consequent SEI layer formation. Unlike in EC:DEC, the use of HC-SC composites exceeds the capacity of their pure components with a maximum of 223mAh/g at 1:1 ratio. However, this capacity is still significantly below any HC-SC composite in EC:DEC, showing that EC:DMC is less suited for use with carbonaceous electrodes other than pure graphite.

As shown in **Fig. 3k)**, the coulombic efficiencies in EC:DMC experience more variation than in EC:DEC. Overall the pure materials have lower 1st cycle efficiencies from 55.2-57.6%. Corresponding to the observed capacities graphite subsequently reaches a high 2nd cycle efficiency of 92.4%, showing that the SEI layer responsible for the low initial efficiency is stable, preventing any further formation in later cycles and protecting the anode. SC on the other has a very low 2nd cycle efficiency of 64.1%, suggesting that the SEI layer on SC is unstable. Among the G-SC composite materials, the 3:1 ratio has a notably high initial efficiency of 67.3% and a reasonably high 2nd cycle efficiency of 90.2%, similar to G-SC 3:1 in EC:DEC. This high 1st cycle efficiency indicates only limited SEI layer formation, and the high 2nd cycle efficiency means that this mild SEI layer is stable. The HC-SC composites show moderate efficiencies, with similar values in the 1st cycle, but quite different magnitudes in the 2nd cycle, reflecting the increased electrolyte stability in the 1:1 ration sample, which is also responsible for the increased capacity.

The rate performance of EC:DMC was found to also be below that of EC:DEC samples. All materials decrease to negligible charge capacities by 0.5C with graphite showing the best recovery again. SC and HC-SC 3:1 show very poor rate capability and reach zero at 0.2C already, while G-SC 3:1 decreases to 34.4mAh/g at 0.2C. Overall, the other composites show similar trends with most recoveries below 50% and related poor capacity retention seen in **Fig. S14**. These results indicate that EC:DMC is a less suitable carbonate-based electrolyte solvent than EC:DEC in all aspects of charge/discharge performance.

3.2.3 Diglyme electrolyte

The GCD curves for the diglyme electrolyte shown in **Fig. S12** have distinctly different shapes to the carbonate-based electrolytes in **Fig. S11**. This is thought to be the result of co-intercalation as the ether-based electrolyte solvates the potassium ions, as observed by Wang[27] for graphite anodes in dimethoxyethane. The effect of using an ether electrolyte is further discussed in 3.3. Agreeing with literature, the depotassiation capacity is lower than in carbonate-based electrolytes. Graphite is found to have the lowest capacity at 133mAh/g, while SC reaches a moderate value of 242mAh/g. Pure HC, however, shows its best performance in ether with a capacity of 265mAh/g. The G-SC composites have an interesting charge curve shape, reaching a plateau at high voltages, the reasons for which should be explored in the future. They show only moderate rate performance of 200-223mAh/g, with a slight maximum at 1:1 ratio, as also observed in EC:DEC. HC-SC composites on the other hand show exceptional capacities of 267-279mAh/g, with a maximum at the 1:1 ratio, similar to EC:DMC. Therefore, the ether-based electrolyte is not particularly suitable for graphite and G-SC, but achieves impressive capacities for HC and HC-SC.

The coulombic efficiencies in **Fig. 3**, show that the different materials experience a variety of efficiencies, especially in the 1st cycle. Graphite exhibits the lowest initial efficiency of 35.3%, suggesting considerable SEI layer formation. While the SC and G-SC composites have moderate 1st cycle capacities between 52.6% and 70.3%, HC and HC-SC composites present remarkably high values of 81.1% for HC and 75.1%, 71.5%, and 77.5% for 3:1, 1:1, and 1:3 HC-SC respectively. However, while the other materials reach above 90% in the 2nd cycle, HC-SC 1:3 only increases to 81.9%, with the other HC-SC composites also experiencing only minor increases in capacity. Compared to the high initial efficiency, this indicates that the SEI layer was not stable and a significant amount of electrolyte decomposes in subsequent cycles.

Disagreeing with the findings of Wang[27], the ether-based electrolyte shows a relatively poor rate performance, as seen in **Fig. 3**. While graphite and G-SC composites recover partially, with G-SC showing a comparatively good ability to store potassium at high rates, HC and HC-SC composites remain at zero capacity when returned to 0.1C. Though potentially related to the cyclability of the material, HC and HC-SC do not show a capacity decay worse than in other electrolytes, such that the reason for the inability to recover must lie elsewhere. As literature[7] suggests that co-intercalation decreases the Stokes radius of the intercalant, such that it could diffuse faster and the anode would undergo a lower volume expansion, which should improve both rate performance and cyclability. Additionally, the HC-SC composites with higher SC ratios undergoes multiple maxima before the rate decays to zero, which has not been observed in any other material. The reasons for these unexpected phenomena are not clear and require further research.

As many interesting trends and anomalies were identified in this overview, it was decided to compare G EC:DEC, G-SC 3:1 EC:DEC, and G ether in more detail. The aim was to gain a deeper understanding of the effect of the electrolyte on graphite intercalation to explain the poor performance in ether, and to understand the high initial coulombic efficiency of G-SC 3:1 in carbonate electrolyte solvents.

3.3 Detailed Study

Fig. 4 contrasts the rate and cycling performance of three different electrode-electrolyte combinations. As discussed in 3.2, graphite in EC:DEC has an excellent capacity at 0.1C, and good rate performance, with exceptional recovery. In comparison, G-SC 3:1 in EC:DEC shows a higher initial capacity but worse rate retention during rate cycling. Graphite in diglyme shows the lowest capacities. All suffer from negligible capacity at higher charge rate of 1 C and 2 C. Graphite in EC:DEC has excellent cycling performance at 0.1C, while G-SC 3:1 in EC:DEC possesses steady cycling ability, with respective capacity retentions of 98% and 73.7% within 17 cycles. However, the capacity of graphite in Ether decays dramatically and the fluctuating efficiency may be a result of electrolyte decomposition or unidentified side reactions.

Further electrochemical characterisation of graphite in both KPF_6 EC: DEC and Ether and G-SC 3:1 in KPF_6

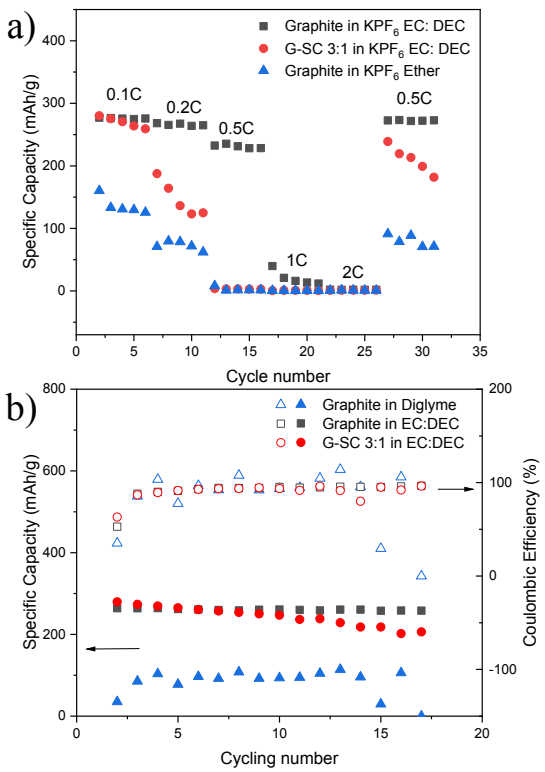


Fig. 4 – a) Rate Performance and b) Cycling Performance comparing graphite in EC:DEC, G-SC 3:1 in EC:DEC, and graphite in Diglyme for capacity and coulombic efficiency

EC:DEC is illustrated in **Fig. 5**. GCD tests were performed to investigate the typical shapes of initial charging and discharging curves, reflecting the initial irreversible capacity loses and working potential for each combination. During 1st cycle potassiation, the initial discharge capacity is much larger than the charge capacity, leading to the low 1st cyclic coulombic efficiency of 35.52%, 54.08% and 62.33% for graphite in Ether, Graphite in EC:DEC and GSC31 in EC:DEC respectively. Comparing the 1st and 2nd cycle of potassiation, the slope region of potassiation shrank from 1.5 V to 0.1 V and little deviation in later cycles. The reduced capacity is attributed to formation of SEI layer with electrolyte decomposition and irreversible intercalation of potassium ion into the anode. The high 1st coulombic efficiency of G-SC 3:1 in KPF6 EC:DEC shows the least irreversible capacity and SEI layer formation. CV further confirms the SEI layer formation and irreversible K-ion intercalation at 1st cycle via the disparity of peak heights and area between the 1st cycle and following cycles.

The curvature of GCD and working potential are quite similar for both graphite and G-SC 3:1 in KPF_6 EC: DEC, which are completely different to that of graphite in Ether, due to the different electrolyte - potassium ion interaction and consequent intercalation mechanisms. The curvature of GCD potassiation into graphite matches the theory of potassium ion intercalation[6], which demonstrates that the changing slopes of one complete discharging curve represents formation of different graphite intercalation compounds (GIC: KC_8, KC_{24} and KC_{36}) in **Fig. S21**. Combined with the CV shape of initial cycling during voltage swings, the working potentials were compared. CV peaks show that the working potential of G-SC 3:1 at around 0.5V is

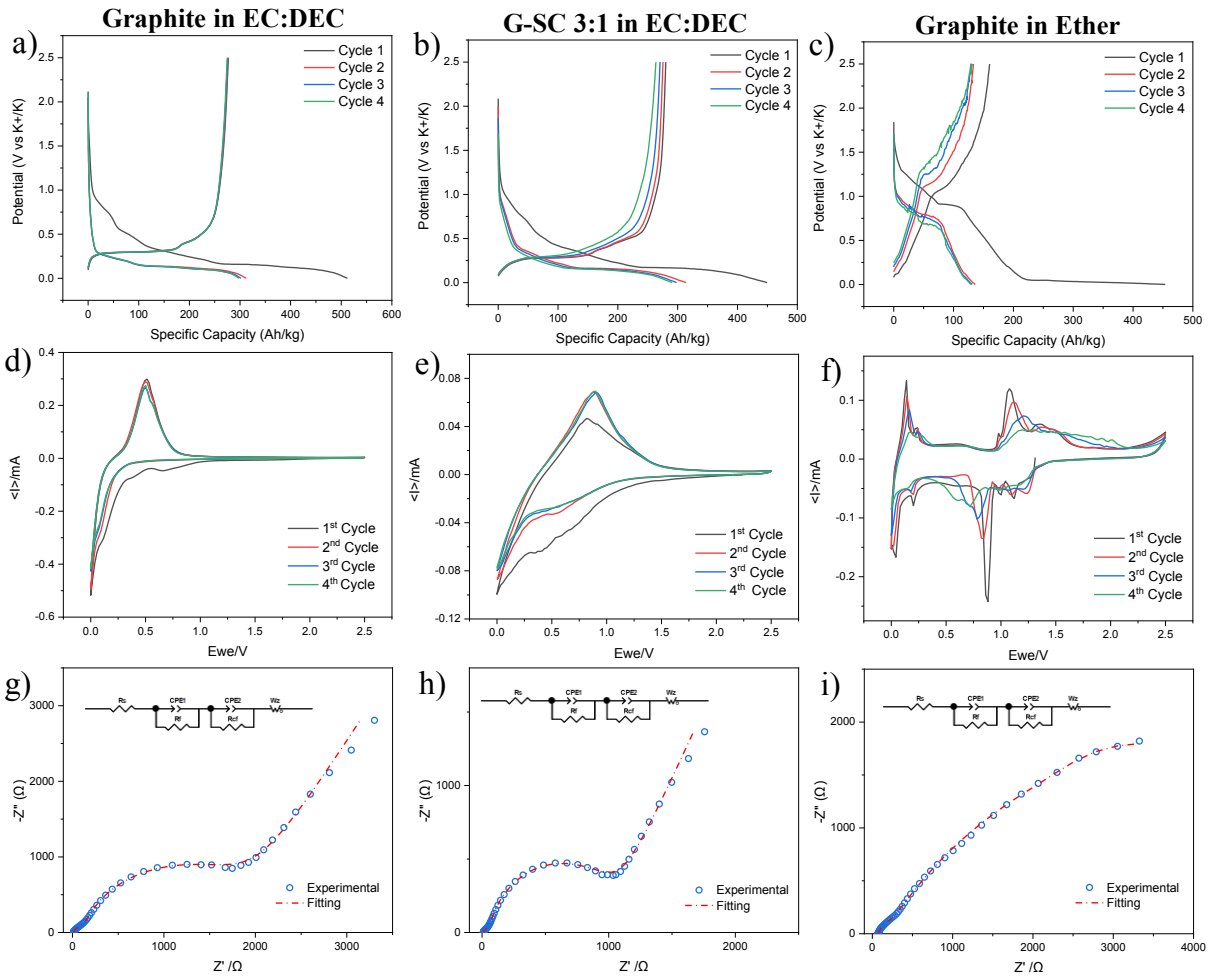


Fig. 5 – GCD profiles at 0.1C of a) graphite in EC:DEC, b) G-SC 3:1 in EC:DEC, and c) graphite in Diglyme; CV for 4 cycles of d) graphite in EC:DEC, e) G-SC 3:1 in EC:DEC, and f) graphite in Diglyme; EIS of g) graphite in EC:DEC, h) G-SC 3:1 in EC:DEC, and i) graphite in Diglyme

higher than that of graphite at 0.15 V vs. K^+/K. Additionally, the peak width demonstrates that the intercalation potential of partially amorphous G-SC 3:1 is broader than that of pure graphite. The composite of soft carbon and graphite may create defects and active sites for favourable potassium storage located at the interface between soft carbon and graphite, leading to an elevated discharge voltage by the synergy between graphite and SC. However, the working voltage of 0.8 V of graphite in ether is significantly higher than in EC:DEC. The same phenomenon was reported [27, 32, 33] for intercalation of Na and K into natural graphite in the presence of $NaPF_6$-Diglyme and KPF_6-DME, respectively. The co-intercalation mechanism for ether-based electrolyte explains the formation of K-ether complex which is subsequently co-inserted into graphite to form a K-ether-Graphite complex. The difference of the Gibbs free energy between pristine graphite and K-ether-Graphite complex leads to the higher standard operating potential according to the Nernst equation ($\Delta G=-nFE$)[32, 33]. Additionally, the different number of ether molecules surrounding the potassium ion influences the working potential. But the fluctuation of GCD at 4th cycle with decreasing peaks in CV substantiate excessive irreversible intercalation and instability. Thus, the formation of K-diglyme-Graphite complex require further investigation of the electrochemical and chemical stability and reversibility of the complex presented in graphite structure. For all material combinations, the voltage plateaus and slopes are situated above the plating potential of potassium metal, avoiding the hazards of dendrite growth.

EIS was conducted to disclose the interface structures at the anode by fitting a typical Nyquist plot for the electrodes after cycling with their respective ending Open Circuit Voltage (OCV). The equivalent circuits are displayed in **Fig. 5**, where R_s (Ω) is the resistance in electrolyte, R_{ct} (Ω) is the charge-transfer resistance, R_f (Ω) is the resistance of the SEI film. Two consecutive semicircles emerge in the high frequency regions (~13k Hz) and moderate frequency region (~100 Hz) corresponding to SEI layer formation and intrinsic charge transfer though the electrode interface, respectively, following the linear region at low frequency (~0.1 Hz) representing the Warburg diffusion of K^+ anions in the bulk carbon[34, 35]. Detailed resistance data and equivalent circuit fitting are **Fig. S22**. The largest charge transfer resistance (3037 Ω) and resistance of SEI layer (119.1 Ω) was found for graphite in Ether, compared to that of graphite in EC:DEC (1600 Ω and 104.9 Ω) and G-Sc 3:1 in EC:DEC (917 Ω and 43.7 Ω), depicting that the ether-based electrolyte results in sluggish charge transfer kinetics and intensive SEI layer formation. The large charge transfer resistance for graphite in Ether may be attributed to the desolvation of the K-ether complex, and ether may be retained between graphite layer and pure solvent without the salt may limit charge transfer conductivity. The lowest resistance

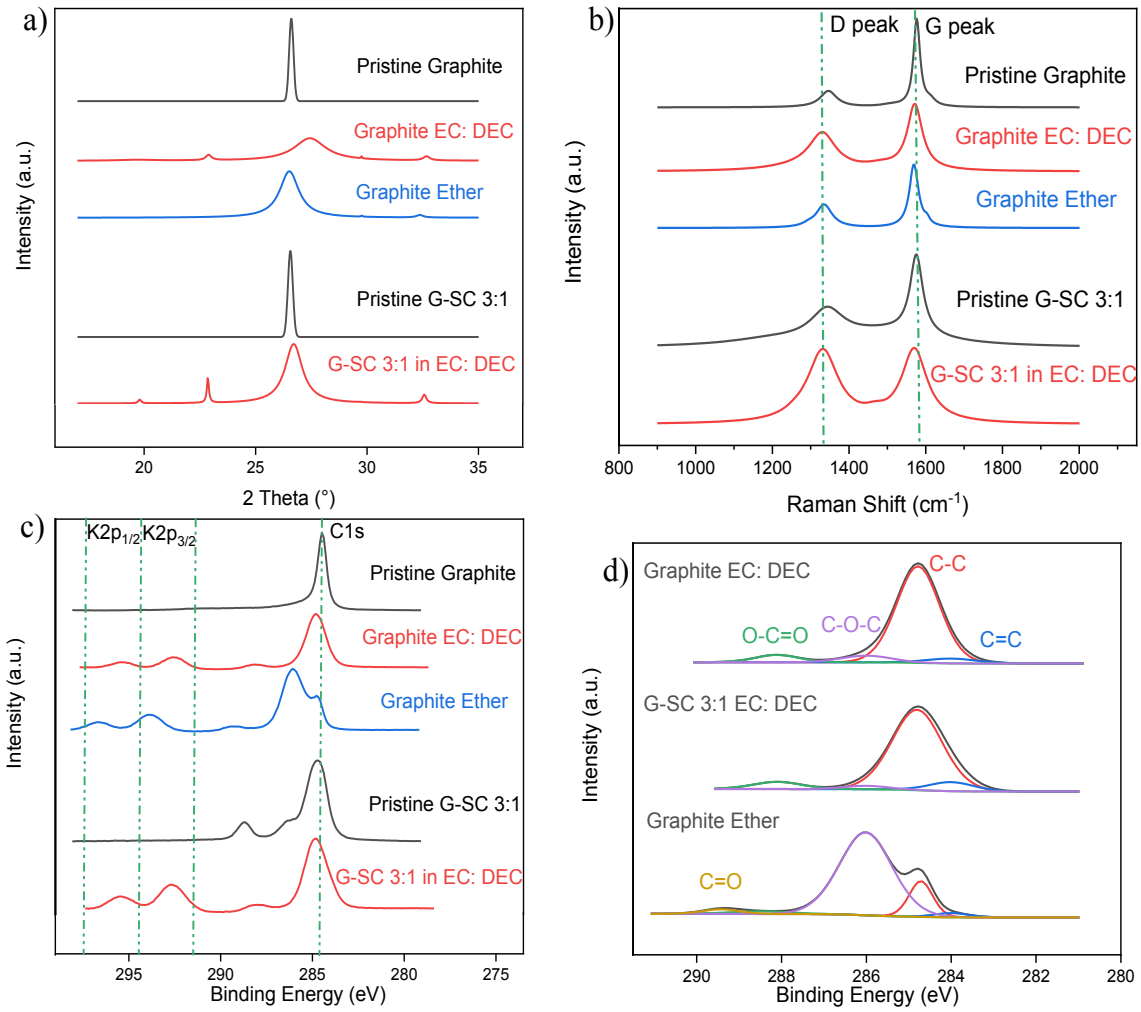


Fig. 6 – Ex-situ a) XRD, b) Raman spectroscopy, c)-d) C1s XPS and C1s XPS fitting of graphite in EC:DEC, G-SC 3:1 in EC:DEC, and graphite in Diglyme in pristine and cycled (discharge-charge-discharge at 0.1C) conditions

for G-SC 3:1 verified its mild SEI layer formation and prove its lowest 1st cycle coulombic efficiency among three combinations.

From BET method in **Fig. S20**, the surface area for graphite and G-SC 3:1 anodes are 21.8 m^2/g and 8.46 m^2/g respectively. The low G-SC 3:1 surface area verifies that fewer sites for SEI layer formation result in its high 1st coulombic efficiency. Furthermore, fast ion transport is supported in both graphite and G-SC 3:1, through their ordered frameworks with hierarchical pores and ion storage in the active sites, including defects and nanopores[36].

The structure variation of the anode materials before and after potassium storage was investigated using XRD and Raman Spectroscopy as shown in **Fig. 6**. The pristine graphite and G-SC 3:1 electrode both show the sharp intrinsic graphite peak at a 2-theta angle of 26.5°. For graphite EC:DEC, the characteristic peaks shift positively to 27.4° with broadened full width at half maximum (FWHM), after full potassiation. According to Bragg's Law, the average spacing between graphitic layer decreases from the intrinsic 3.35 Å to 3.24 Å. The peak shift is also observed in Jian's work[6] in which shows as shoulder peaks around 27.5°. The intercalation of potassium into graphite with EC:DEC may expand and tilt some layer of graphite structure and compress the nearby layer into smaller spacing, leading to shift of the characteristic peak and formation of shoulder peaks. **Fig. 7** compares the non-ideal intercalation with ordered allocation of potassium ions and parallel graphite layers, where the number of graphene layers between potassium varies with GIC stage[6]. The shift of the 002 peak is inhibited in the case of G-SC 3:1 in EC:DEC, which illustrates that soft carbon has a protecting effect to maintain the mechanic structure of pure graphite layer and avoid the inclination of graphite layers. For both graphite and G-SC 3:1 in EC:DEC, pairs of small peaks are located at 19.8°& 32.6°, and 22.8° & 29.7° which are representative peaks for GICs KC_{24} and KC_{36} respectively. After fully potassiating the anode, the KC_{24} and KC_{36} are not completely converted into KC_8, which prevents it from reaching theoretical capacity. The accumulation of GICs also harms the cycling performance and graphite structure. For diglyme, the 002 peak remains at 26.5°. The potassium ion is solvated

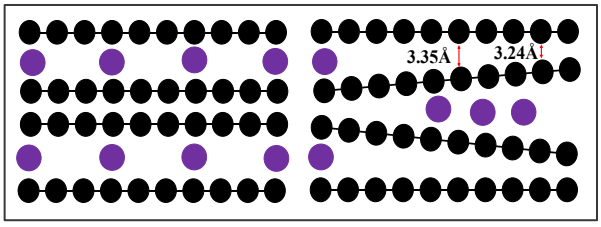


Fig. 7 – ideal and proposed non-ideal intercalation of potassium (purple) into graphite (black) in EC-DEC

by diglyme electrolyte with a smaller Stokes radius, which is considered to reduce the volume expansion of graphite, aiming to improve the cycling ability[7, 11]. The co-intercalation mechanism in ether is understood to be different to the intercalation process of the carbonate. However, the XRD of graphite in diglyme exhibits small peaks at 29.7° and 32.6° corresponding to intercalation. It suggests that the K-Diglyme complex experiences desolvation in graphite layer and forms GICs instead of K-Diglyme-Graphite complexes insertion into the graphite layer, while diglyme may also diffuse into the anode solely. The trapped diglyme in the structure of graphite may be responsible for the continuous capacity decay and fluctuation GCD curve of decomposition. The broadened peaks (002) with lower intensity after potassiation suggestions a certain extent of deformation to graphite for all three combinations.

The fittings of Raman Spectroscopy in **Fig. S19** for I_D/I_G can be found in Supporting material. The original I_D/I_G intensity ratio between D peak and G peak is 0.187 and 0.403 for pristine graphite and pristine G-SC 3:1, with PVDF and carbon black. After potassium ion insertion, the I_D/I_G intensity ratio enlarge to 0.374 for graphite in Diglyme, 0.546 for graphite in EC: DEC and 1.00 for G-SC 3:1 in EC: DEC. A phenomenological Three-Stage model[28-30] is used as an indicator to interpolate the meaning of I_D/I_G for different case. Pristine graphite and pristine G-SC 3:1 are both locate in stage 1, as explained in 3.1. After cycling, the 2D peaks for graphite in diglyme, graphite in EC:DEC and G-SC 3:1 in EC:DEC remain sharp, proving all combinations remain in stage 1, where the density of defects and sp^3 content increase with the increase of the I_D/I_G ratio. The I_D/I_G ratio of graphite in EC:DEC is higher than that in Diglyme, which infers that the formation of sp^3-defects and structural damage to graphitic layer is more severe in EC:DEC. The evaluation from Raman spectroscopy is consistent with that from XRD, showing less structural damage by the diglyme electrolyte. After full potassiation, the I_D/I_G ratio of G-SC 3:1 in EC:DEC is boosted. The percentage increase of I_D/I_G for G-SC 3:1 is smaller than that for graphite in EC: DEC, which implies a more ordered structure in G-SC 3:1 and confirms the effective protection offered by introduction of SC to the graphitic structure. Ex-situ XRD and Raman have shown that the synergy of graphite and soft carbon results in mechanically stability in EC:DEC. The intercalation mechanism for diglyme electrolyte not only suffers from volume expansion by forming GICs between graphite layer but also the structural integrity is further hampered by the mechanism of co-intercalation.

Other than evaluating the deformation of graphite layers, XPS in **Fig. 6** was conducted to investigate the SEI layer contents on the anode surface after potassiation. Two peaks of C1s and O1s for pristine materials and four distinct peaks were observed after cycling, corresponding to C1s, K2p, K2s and O1s for all combinations of anode and electrolytes in full range XPS **Fig. S17**. The inconspicuous peak (<2%) around 688.4 eV is F1s resultant from the PVDF binder and the KPF_6 salt. The ratios between O1s and C1s are interpolated from the intensity of corresponding peaks for each combination to depict the decomposition of both carbonate-based and ether-based electrolytes. The two minor peaks around 280 eV for ex-site characterisation affirm the composition of $K2P_{1/2}$ and $K2P_{3/2}$ which may be the intercalation or irreversible deposition on the surface of anode. Further XPS would be required to detect the potassium peak after depotassiation to examine the state of potassium on the anode surface. The O1s/C1s ratio of pristine graphite is 3.85 % which significantly increases to 15.62% and 13.91% for Graphite in EC:DEC and in diglyme. Although the ratio for graphite in diglyme is slightly lower than that in EC:DEC, the diglyme may also decompose inside graphite structure instead of remaining at surface. After potassiation, the O1s/C1s ratio of G-SC 3:1 in EC: DEC is 20.18% is unexpectedly diminished from the O1s/C1s ratio of 23.14% for pristine G-SC 3:1. The major oxygen content in the pristine material originates from the oxygen content in pitch. The drop in oxygen content after cycling requires further investigation. It is believed that the originally high composition of oxygen may passivate the interface between electrolyte and electrode in advance and inhibit SEI layer formation. Additionally, the oxygen content enhances the wettability of the anode by the electrolyte, leading to faster charge transfer kinetics. The decreased O1s/C1s ratio may also be closely related to the formation mechanism of the thin SEI layer and highest 1st coulombic efficiency for G-SC 3:1 composite. O1s scan for post-potassiation samples are further fitted with different bonds. There is clear difference between the two different electrolytes. While they all possess C-C and C=C bonds, EC:DEC has a higher peak intensity for O-C=O and relatively small amount of C-O-C regardless of anode material, which is the decomposition product from EC:DEC. However, the SEI content has large amount of C-O-C bond and C=O bond in the case for diglyme. It is presumed that the C-O-C bond from due to ether decomposition. The large intensity peak of C-O-C and relatively smaller peak for C=O is consistent with the result from electrochemical characterisation of high-resistance SEI layer formation, limiting graphite performance in ether.

4. Conclusion

The synergic effect between graphite and soft carbon, contributing to the excellent performance of G-SC composite, is confirmed. The amorphous structure of soft carbon provides adsorption and intercalation sites for potassium ion insertion which elevates the working potential. Also, the addition of soft carbon provides protection and flexibility to the graphitic layers, resulting in minor structural damage. The unexpected high oxygen content in pristine G-SC 3:1 composite may passivate the material aforehand inhibiting the SEI layer formation. Another important insight of this research is that the intercalation mechanism behaves differently in the presence of carbonate-based electrolyte and ether-based electrolyte. For diglyme, it exhibits both intercalation mechanism and co-intercalation. The entangled mechanism complicates the potassium ion insertion into graphite, resulting in intensive SEI layer formation and extremely unstable cycling performance. It is possible that the solvation of

potassium ion in different ether affects the K storage mechanism.

The overview of multiple anode and electrolyte materials has shown, that we could not confirm other researchers' findings of superior performance in SC and HC compared to graphite, with this traditionally versatile material proving particularly reliable in carbonate-based electrolyte. G-SC was found to supersede the electrochemical performance of its pure components in some respects in carbonate-based electrolytes, while HC and HC-SC composites appear promising and intriguing in the ether-based electrolyte. The composites show ratio-dependent performance maxima, proving that composition tuning is key to further develop batteries in the future.

Acknowledgements

We would like to thank the PhD students Zhenyu (Johny) Guo and Zhen Xu for their continuous guidance.

References

[1]Dunn, B., H. Kamath, and J.M. Tarascon, *Electrical Energy Storage for the Grid: A Battery of Choices.* Science, 2011. **334**(6058): p. 928-935.

[2]Li, Y., et al., *Intercalation chemistry of graphite: alkali metal ions and beyond.* Chemical Society Reviews, 2019. **48**(17): p. 4655-4687.

[3]Pramudita, J.C., et al., *An Initial Review of the Status of Electrode Materials for Potassium-Ion Batteries.* Advanced Energy Materials, 2017. **7**(24).

[4]Tarascon, J.M., *Is lithium the new gold?* Nature Chemistry, 2010. **2**(6): p. 510-510.

[5]Kuksenko, S.P., *Aluminum foil as anode material of lithium-ion batteries: Effect of electrolyte compositions on cycling parameters.* Russian Journal of Electrochemistry, 2013. **49**(1): p. 67-75.

[6]Jian, Z.L., W. Luo, and X.L. Ji, *Carbon Electrodes for K-Ion Batteries.* Journal of the American Chemical Society, 2015. **137**(36): p. 11566-11569.

[7]Matsuda, Y., et al., *Behavior of Some Ions in Mixed Organic Electrolytes of High-Energy Density Batteries.* Journal of the Electrochemical Society, 1981. **128**(12): p. 2552-2556.

[8]Zhao, Q., et al., *Oxocarbon Salts for Fast Rechargeable Batteries.* Angewandte Chemie-International Edition, 2016. **55**(40): p. 12528-12532.

[9]Xue, L.G., et al., *Liquid K-Na Alloy Anode Enables Dendrite-Free Potassium Batteries.* Advanced Materials, 2016. **28**(43): p. 9608-+.

[10]Xing, Z.Y., et al., *A perylene anhydride crystal as a reversible electrode for K-ion batteries.* Energy Storage Materials, 2016. **2**: p. 63-68.

[11]Wang, X.P., et al., *Earth Abundant Fe/Mn-Based Layered Oxide Interconnected Nanowires for Advanced K-Ion Full Batteries.* Nano Letters, 2017. **17**(1): p. 544-550.

[12]Jian, Z.L., et al., *Poly(anthraquinonyl sulfide) cathode for potassium-ion batteries.* Electrochemistry Communications, 2016. **71**: p. 5-8.

[13]Eftekhari, A., *Potassium secondary cell based on Prussian blue cathode (vol 126, pg 221, 2004).* Journal of Power Sources, 2004. **136**(1): p. 201-201.

[14]Chen, Y.N., et al., *Organic electrode for non-aqueous potassium-ion batteries.* Nano Energy, 2015. **18**: p. 205-211.

[15]Jian, Z.L., et al., *Hard-Soft Composite Carbon as a Long-Cycling and High-Rate Anode for Potassium-Ion Batteries.* Advanced Functional Materials, 2017. **27**(26).

[16]Jian, Z.L., et al., *Hard Carbon Microspheres: Potassium-Ion Anode Versus Sodium-Ion Anode.* Advanced Energy Materials, 2016. **6**(3).

[17]Vaalma, C., et al., *Non-Aqueous K-Ion Battery Based on Layered K0.3MnO2 and Hard Carbon/Carbon Black.* Journal of the Electrochemical Society, 2016. **163**(7): p. A1295 A1299.

[18]Sultana, I., et al., *Tin-based composite anodes for potassium-ion batteries.* Chemical Communications, 2016. **52**(59): p. 9279-9282.

[19]Han, J., et al., *Exploration of K2Ti8O17 as an anode material for potassium-ion batteries.* Chemical Communications, 2016. **52**(75): p. 11274-11276.

[20]McCulloch, W.D., et al., *Potassium-Ion Oxygen Battery Based on a High Capacity Antimony Anode.* Acs Applied Materials & Interfaces, 2015. **7**(47): p. 26158-26166.

[21]Fan, L., et al., *Graphite Anode for a Potassium-Ion Battery with Unprecedented Performance* Angewandte Chemie-International Edition, 2019. **58**(31): p. 10500-10505.

[22]Cao, B., et al., *Mesoporous soft carbon as an anode material for sodium ion batteries with superior rate and cycling performance.* Journal of Materials Chemistry A, 2016. **4**(17): p. 6472-6478.

[23]Winter, M., et al., *Insertion electrode materials for rechargeable lithium batteries.* Advanced Materials, 1998. **10**(10): p. 725-763.

[24]Perner, A. and J. Vetter, *Lithium-ion batteries for hybrid electric vehicles and battery electric vehicles*, in *Advances in Battery Technologies for Electric Vehicles*. 2015, Woodhead Publishing: Cambridge. p. 173-190.

[25]Khosravi, M., N. Bashirpour, and F. Nematpour, *Synthesis of Hard Carbon as Anode Material for Lithium Ion Battery.* Advanced Materials Research, 2013. **829**: p. 922-926.

[26]Xie, F., et al., *Hard-Soft Carbon Composite Anodes with Synergistic Sodium Storage Performance.* Advanced Functional Materials, 2019. **29**(24).

[27]Wang, L.P., et al., *Graphite as a potassium ion battery anode in carbonate-based electrolyte and ether-based electrolyte.* Journal of Power Sources, 2019. **409**: p. 24-30.

[28]Ferrari, A.C. and D.M. Basko, *Raman spectroscopy as a versatile tool for studying the properties of graphene.* Nature Nanotechnology, 2013. **8**(4): p. 235-246.

[29]Ferrari, A.C., et al., *Raman spectrum of graphene and graphene layers.* Physical Review Letters, 2006. **97**(18).

[30]Ferrari, A.C. and J. Robertson, *Interpretation of Raman spectra of disordered and amorphous carbon.* Physical Review B, 2000. **61**(20): p. 14095-14107.

[31]Xu, Z., et al., *All-Cellulose-Based Quasi-Solid-State Sodium-Ion Hybrid Capacitors Enabled by Structural Hierarchy.* Advanced Functional Materials, 2019. **29**(39).

[32]Kim, H., et al., *Sodium Storage Behavior in Natural Graphite using Ether-based Electrolyte Systems.* Advanced Functional Materials, 2015. **25**(4): p. 534-541.

[33]Jache, B. and P. Adelhelm, *Use of Graphite as a Highly Reversible Electrode with Superior Cycle Life for Sodium-Ion Batteries by Making Use of Co-Intercalation Phenomena.* Angewandte Chemie-International Edition, 2014. **53**(38): p. 10169-10173.

[34]Xiao, L.F., et al., *Hard carbon nanoparticles as high-capacity, high-stability anodic materials for Na-ion batteries.* Nano Energy, 2016. **19**: p. 279-288.

[35]Prabakar, S.J.R., J. Jeong, and M. Pyo, *Nanoporous hard carbon anodes for improved electrochemical performance in sodium ion batteries.* Electrochimica Acta, 2015. **161**: p. 23-31.

[36]Trogadas, P., et al., *A lung-inspired approach to scalable and robust fuel cell design.* Energy & Environmental Science, 2018. **11**(1): p. 136-143.

Effect of high energy facets on the crystal size distributions of seeded batch processes

FERDINAND AGU AND ISHA BADE

Department of Chemical Engineering, Imperial College London. UK
Compiled December 12, 2019

Abstract

In pharmaceutical crystallisation it is desirable to produce a final product with a mono-dispersed Crystal Size Distribution (CSD) and large median size, which is rather challenging due to the stochastic nature of crystallisation. A major contributor to this stochasticity is the effect of breakage on crystals arising from upstream processes; milling and filtering of seeds, as well as crystal-crystal and crystal-impeller interactions in the crystalliser. These mechanisms are poorly understood and are not thoroughly accounted for in industry standard process modelling tools such as gPROMS®. This work details the engineering approach of producing paracetamol crystals with the aforementioned characteristics from a paracetamol-ethanol system using lab-scale batch crystallisation and concurrently modelling the process on gPROMS® to investigate the limits of the software. Experimental results are compared to the model simulation and based on the observations made, the behaviour of broken fragments formed by secondary nucleation via attrition is studied. An In-depth description of the DSLR camera set-up and the Python™ code used to capture and analyse this behaviour on a single crystal is outlined in the paper. Consequently, a mechanism of 'self-healing' on broken crystals with high energy facets has been proposed which explains the behaviour of broken crystals in a batch scale operation . This mechanism is in reasonable agreement with the mathematical model developed by Mersmann for the growth of attrition fragments and the following paper highlights the need for it to be implemented in modelling tools to increase the accuracy of industrial modelling.

1. INTRODUCTION

A number of catastrophic incidents resulting from the inefficiencies in downstream processing, such as the thalidomide crisis – in which the (R)-(+)- thalidomide isomer was used predominantly as a sedative by pregnant women; whereas, its optical isomer (S)-(-)- thalidomide was unfortunately found to cause malformation and ultimately the death of thousands of infants; have emphasised the urgency for a consistent means of producing active pharmaceutical ingredients (APIs) that meets the stringent demands of the pharmaceutical industry; including chirality, product purity, crystal form and crystal size distribution (CSD)(Gao et al., 2017). Consequently, crystallisation as an industrial technique has rapidly gained traction over alternative methods of producing crystals due to its high selectivity and efficiency, as well as the benefit of separation, purifying and refining of APIs in a single process step (Artusio and Pisano, 2018).

Currently, most pharmaceutical crystallisation is carried out by seeding as this method enables better control and reproducibility of the process; and in batches, to allow for the production of various APIs with different design requirement, as well as a changing market. However, these processes are difficult to model and control accurately as they tend to be stochastic, yielding batch-to-batch variations in the final product specification(Su, Nagy, and Rielly, 2015). The production of a desired API typically involves a number of upstream stages for seed preparation such as drying, milling, and filtering. These processes typically exert mechanical stress on the seeds, which leads to breakage – therein lies a significant contributor to the stochasticity of the batch crystallisation process, the effect of breakage.

The "Made Smarter. Review 2017" highlights the need for innovative ways of developing manufacturing processes based on digital design techniques and mathematical models in an attempt to cut costs (UK Government, 2017). gCRYSTAL® is the state-of-the-art crystalli-

sation library of industry standard modelling software – gPROMS®. It utilizes kinetic models developed to account for the evolution of particles in the various phases of a crystal's life cycle, together with a population balance model to simulate the process. The phases accounted for in the model include - nucleation (activated or via attrition), growth and agglomeration. Although these models often produce substantial estimates for initial modelling; they are limited as they do not rigorously account for the effect breakages within the crystalliser as a result of crystal-crystal and crystal-impeller interactions. This is highlighted by the notable discrepancy between the crystal size distributions (CSD) of experimental crystallisation and the software simulated output.

The crystallisation process in this paper uses a paracetamol and ethanol system, due to the prevalence of the former in the pharmaceutical industry. It was engineered to produce a mono-dispersed final size distribution (FSD) with large median size crystals, as these are desirable in industry to increase the efficiency of downstream processes (Noriaki Kubota et al., 2001).

Drawing upon the above, the effect of breakage on the CSD was investigated via a series of laboratory scale batch crystallisation processes. These experiments were then simulated using gCRYSTAL® and the resultant CSDs were compared to provide insight into the limitations of the software. A self-healing hypothesis was subsequently investigated to account for the observed discrepancy.

2. THEORY

Crystallisation is a two-step process initiated by nucleation - advent of a new phase, followed by the growth of a critical i.e stable nuclei into crystals of visible size (Mullin, 2001). This section provides a deeper insight into the theories of nucleation and growth along with the kinetic models currently used in gPROMS® to quantify these theories. It also provides detailed overview of the fundamental principles of crystallisation; which in conjunction with the nucleation and growth theories provides a basis for predicting the desired product properties.

2.1. Nucleation

Nucleation can be broken down into two main categories; 'primary' where nucleation occurs in the absence of crystalline matter solely due to the solution reaching supersaturation, and 'secondary' where crystals present in the solution are used as a site and source for the formation of nuclei (Agrawal and Paterson, 2015).

Since secondary nucleation is the dominant nucleation mechanism in industrial crystallizers, the understanding of the growth behaviour of the secondary nuclei is critical for predicting and controlling the crystallization process (Virone et al., 2005). Extensive literature suggests the existence of various mechanisms of secondary nucleation such as contact (shear) nucleation, attrition, activated nucleation, initial/surface breeding and seeding (Mullin, 2001),(Agrawal and Paterson, 2015). As this paper focuses mainly towards seeding and breakage as well as touches upon activated secondary nucleation, the following sections explain them in relative depth.

2.1.1. Seeding

Most of the secondary nucleation mechanisms are unintentional and are therefore difficult to control. Seeding on the other hand, which is a deliberate attempt at introducing crystalline nuclei in the system, is frequently employed in industrial crystallisation to effect control over the product size and size distribution (Mullin, 2001) along with enabling operations to be kept at a lower supersaturation throughout the batch if enough seeds are loaded (Kobari, N. Kubota, and Hirasawa, 2010).

2.1.2. Activated nucleation

This is a two step process where the first part involves the activation required to form secondary nuclei on the surface of a crystal followed by the second step where these nuclei detach from the surface and consequently grow (Verdurand et al., 2005). Activated nucleation can be difficult to control as it has non-linear kinetics (Verdurand et al., 2005) and is unlikely to happen in industrial crystallizers due to the requirements of high supersaturation (Myerson, 2001).

2.1.3. Attrition nucleation

Attrition, a mechanical process, is independent of supersaturation, and refers to removal of fragments from the parent crystal due to collisions with each other (crystal-crystal interactions) or some other objects (crystal-impeller interactions). The generated fragments act as nuclei and are likely to be a major source of secondary nucleation in industrial crystallisers (Agrawal and Paterson, 2015).

The Evans kinetic model is the recommended model in gCRYSTAL® for the secondary nucleation arising from attrition. It assumes the rate of secondary nucleation to be proportional to the product of collision frequency and impact energy arising from crystal-crystal (CC) and crystal-impeller (CI) interactions (Evans and Sarofim, 1974).

The Evan's rate expression for CC and CI interactions are shown in equations 1 and 2 respectively where k_n is the rate constant for secondary nucleation, n is the order of supersaturation dependency and L_{min} is the size above which crystals undergo attrition.

$$J_{sec_ci} = \exp(\ln k_{n_ci})\sigma^{n_{ci}} \frac{N_Q}{N_P} k_v \rho_c \epsilon \int_{L_{min_ci}}^{\infty} nL^3 dL \quad \textbf{(1)}$$

$$J_{sec_cc} = \exp(\ln k_{n_cc})\sigma^{n_{cc}}\rho_c\epsilon^{5/4}L^{-4}\int_{L_{min_cc}}^{\infty} nL^0 dL \quad \textbf{(2)}$$

2.2. Growth

The behaviour of the growth rate of a crystal face is dependent on its underlying growth mechanism (L. Li and Rodríguez-Hornedo, 1992) which ultimately decides the crystal size. Surface energy theories for growth propose that crystals will grow to minimise their surface free energies to gain the most stable shape, but over time this approach has been receiving reduced acceptance (Mullin, 2001). Diffusion theories, presume that growth is a continuous process where matter is constantly deposited on the crystal surface at a rate proportional to the difference in concentration between the crystal surface and the bulk of the solution (Mullin, 2001). However, Volmer (1939) suggested that crystal growth was a discontinuation process, taking place by adsorption, layer by layer, on the crystal surface and this led to the adsorption layer theories, several notable modifications of which have been proposed in recent years (Mullin, 2001). There are three generally accepted mechanisms of the above mentioned kind of crystal growth, each of which reflects a particular range of supersaturation; screw dislocation (BCF) mechanism, Birth & Spread (B&S) meachanism and rough interfacial growth model with increasing supersaturation. (Nguyen et al., 2014).

These growth mechanisms are achieved by a 2-step kinetic model developed by Mersmann (Myerson, 2001), where diffusion of solute molecule to the crystal surface is followed by its integration in the lattice. gPROMS® uses a similar kinetic model to simulate these growth theories except in the software, the relative supersaturation $\left(\frac{C_{int}(L)-C_{sat}}{C_{sat}}\right)$ is employed in the surface integration expression, instead of the ΔC , $(C_{bulk} - C_{sat})$, between the bulk solution and the crystal interface (PSE Model Documentation, 2015). The mass transfer step for a diffusion-controlled crystal growth rate (due to the low supersaturation) is expressed by the following equation:

$$G(L) = k_d(L)\left[\frac{C_{bulk} - C_{int}(L)}{\rho_{crys}}\right] \quad \textbf{(3)}$$

The expression for surface integration (equation 4) accounts for the above diffusion step, orientation of the solute molecule in order to slot into the crystal lattice, and the actual integration into the lattice.

$$G(L) = k_g \exp\left(\frac{-E_{A,g}}{RT}\right)\left[\frac{C_{int}(L) - C_{sat}}{C_{sat}}\right]^g \quad \textbf{(4)}$$

Where k_g, g, $E_{A,g}$ are the growth rate constant, order of supersaturation dependency and activation energy respectively and applies if $C_{bulk} - C_{sat} > 0$

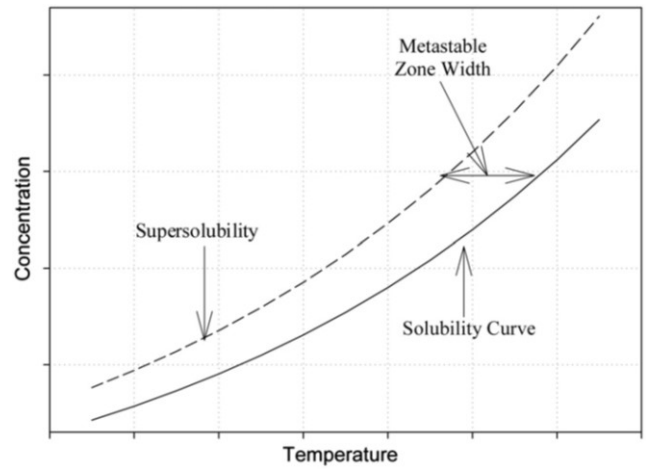


Fig. 1. Solubility curve showing the MSZ for a common API (Frawley et al., 2012)

2.3. Crytallisation process

A graphical representation of the thermodynamic phases involved in crystallisation process can be seen in figure 1. The solubility line depicts the equilibrium concentrations of a solution at a given temperature where the solution is regarded as saturated. Below this line, lies the undersaturated region where no crystallisation takes place, while the region above the supersolubility line corresponds to the labile region where spontaneous nucleation occurs. The region between the supersolubility and the solubility curve is known as the metastable zone, the region considered ideal for crystal growth (Parsons, Black, and Colling, 2003). Seeds are introduced in this region to provide optimum control on the growth of the seeds and eventually on the final crystal size.

3. BATCH CRYSTALLISATION AND MODELLING

3.1. Batch Crystallisation Methodology

The seeded laboratory scale batch experiments were engineered from first principle understanding in an attempt to produce crystals with large median size and a unimodal distribution. The conditions used for the process were obtained from the work done by Mitchell and Frawley (2012), shown in figure 2. As such, seeds were introduced under metastable conditions; represented as point '1' in figure 2 , to promote growth and limit the excessive formation of nuclei which could lead to a polydisperse distribution. (Artusio and Pisano, 2018).

The induction time was first investigated via primary homogenous nucleation for a 0.2g/ml concentration of paracetamol in ethanol at a temperature of 20°C, to provide empirical justification for seed timing. Sub-

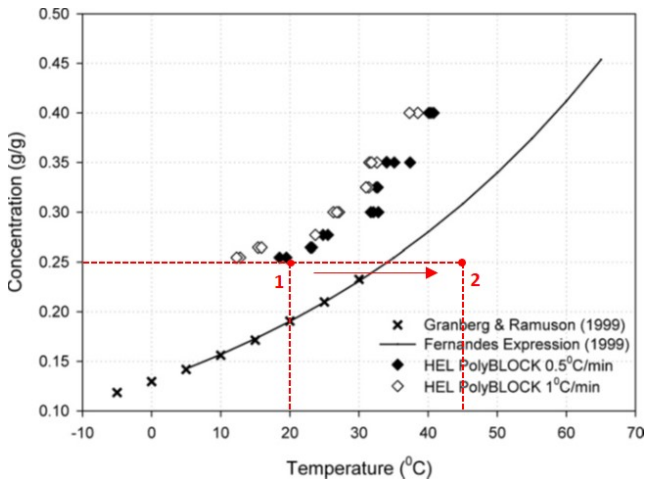


Fig. 2. Literature solubility data (Ramuson, 1999 and Fernandes expression, 1999) and supersolubility curve for a cooling rate of 1°C/min (HEL Polyblock 1°C/min) for a paracetamol-ethanol system (Frawley et al., 2012)

sequently, two seeded runs - one with milled and the other with un-milled seeds, were carried out with two repeats.

3.1.1. Materials

The experimental work described in the following sections was performed using Acetaminophen (Paracetamol) of powdered form and Ethanol, absolute, both sourced from Sigma Aldrich. The substances employed were of analytical standard, used as received and had a purity of $\geq$ 98%. Ethyl acetate, anhydrous, utilised for the CSD measurements in section 3.1.5 was also acquired from Sigma Aldrich at a purity of $\geq$ 98%.

3.1.2. Induction time

A 50 ml CSTR crytstalliser operated at 400 rpm impeller speed was used to determine the induction time. The procedure is depicted by the graph in fig 3 and was adapted from recent literature (H. Li et al., 2017). 50 ml of ethanol was added to 10 g of paracetamol in the crystalliser at room temperature, the solution was then heated to 45°C and held at this temperature for 1 hour to ensure the complete dissolution of paracetamol; temperature determined from point '2' in figure 2 (without the seeding step). A cooling ramp was then set at a cooling rate of 1°C/min until a temperature of 20°C was attained. A Nikon D90 camera was then set up to capture images of the crystalliser every minute to determine the induction time - time taken for the solution to become turbid i.e commencement of primary nucleation. This was repeated 3 times to get an average induction time to account for the stochastic nature of crystallisation.

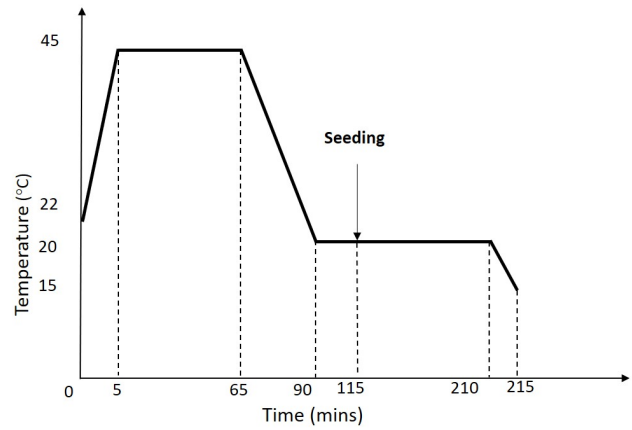


Fig. 3. Temperature Profile for Batch Crystallisation Process

3.1.3. Seed Preparation

The seeds used for the unmilled run were obtained via vacuum filteration of the product from the homogenous crystallisation described in section 3.1.2. It was then left in an incubator at 20°C for 24 hours, after which is was sieved to get a seed size of 64-75µm.

In the case of the milled seeds, a saturated solution of 0.23g/ml of paracetamol was made and left to settle. The supernatant was then pipetted into a 20ml vile which was subsequently covered with parafilm tape with a few holes pierced on top to control the evaporation. The solution was then left at room temperature to evaporate and the resultant seeds were retrieved, crushed in a mortar and sieved to a size of 64-75µm.

3.1.4. Seeded Crystallisation

The seeded crystallisation processes were carried out in accordance with the same temperature profile as in fig 3. 1ml of solution in the crystalliser was taken out as soon as the crystalliser reaches 20°C. This was then added to 1g of seed after which the seeds were pipetted into the crystalliser at 115 minutes into the process. The crystal products were obtained via vacuum filtration and analsysed using the MasterSizer 2000 to obtain the CSD.

3.1.5. CSD Measurement

The CSDs for the products obtained was generated using the MasterSizer 2000, which is operated using the principles of static light scattering and the Mie scattering theory to generate the distribution of particle sizes (CRODA, n.d.). A polar aprotic solvent - ethyl acetate was used as the carrier solvent due to the low solubility of paracetamol in it (10.73g/kg) (Granberg and Rasmuson, 1999), efficient dispersion of solute, and the low propensity for agglomeration to occur.

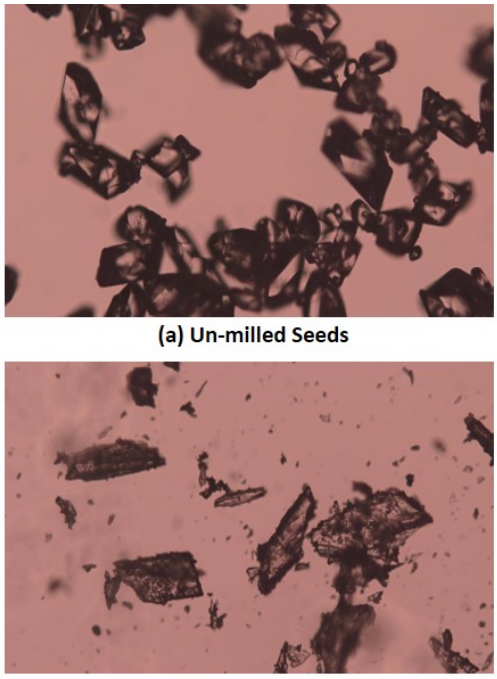


Fig. 4. (a) Un-milled and (b) Milled seeds under a microscope at x20 magnification

3.2. gCRYSTAL® Simulation

The batch crystallisation process covered in the above sections was subsequently modelled on gPROMS® using the gCRYSTAL® library, a very novel approach at simulating batch crystallisation. The physical parameters and conditions implemented in the simulation were as close to the experimental procedure as possible.

3.2.1. Physical Properties Data

The physical properties for Ethanol and Paracetamol were obtained from the DIPPR® and PubChem® database whereas empirical data for the thermodynamic properties of the ethanol-paracetamol mixture were found using literature (Khabriev et al., 2017). Furthermore, the crystallisation process was modelled using a Mixed-Suspension Mixed-Product Remover (MSMPR) crystalliser with a blank inlet and outlet to operate under batch conditions. This category of crystallisers is the most important to chemical process industries and is used when crystal growth is desired as a suspension in solution (BrandConnect, 2010) as was the case at hand. A temperature sensor and controller were used to execute the temperature profile shown in figure 3 along with a particle size distribution sensor to measure the final crystal size. Energy consumption of the system was also accounted for by considering heat transfer between the mixture and the PYREX® crystalliser; cross sectional area of the crystalliser was calculated as that of a top-open cylinder.

3.2.2. Assumptions and Limitations

As with any other modelling approach certain assumptions were made to aid the performance of the simulation; firstly, no primary nucleation was taken into consideration because, as mentioned in section 3.1.4, seeds were introduced under the metastable conditions before any spontaneous nucleation could occur. With a supersaturation ratio of merely 1.3, secondary activated nucleation is unlikely to happen (section 2.1.2) and was therefore disregarded in the modelling scenario. Moreover, recent studies (Simone, 2018) show that using specific temperature cycling i.e a temperature profile similar to the one used in this situation, promotes deagglomeration and hence, no agglomeration has been accounted for in the model. Lastly, the seed and crystal geometry were assumed to be spherical where the milled seeds were defined to be equivalent to a quarter of a sphere using a corresponding shape factor.
The most crucial limitations of the model were found to be the inability to specify the surface chemistry of seeds and the seed timing, gCRYSTAL makes a presumption that seeds are present in the system from the beginning of the process.

3.2.3. Kinetic Models

After taking into consideration the assumptions and limitations, two major kinetic models were implemented in the simulation, one for secondary nucleation via attrition and the other one for growth of crystals. The Evans kinetic model described in section 2.1.3 was used for secondary nucleation arising from attrition and as both the types of interactions were considered in this case, the resulting overall secondary nucleation rate was simply the summation of the expressions for CC and CI collisions (PSE Model Documentation, 2015). Three parameters: (1) Rate constant for secondary nucleation k_n (2) Order of supersaturation dependency n and (3) Size above which crystals undergo attrition L_{min}, were needed to be specified in gPROMS. Due to the lack of adequate literature in this context as well as the inability to carry out experiments to get empirical data, parameters were estimated in relation to the characteristics of the experiment. However, it was made sure that the rate of CC interactions was relatively lower than the rate of CI interactions as suggested by Mullin (Mullin, 2001).
The growth and dissolution model recommended by gPROMS is detailed in section 2.2 and was employed for the simulation. Using the equations 3 and 4 three parameters: (1) Growth rate constant k_g (2) Order of supersaturation dependency g and (3) Activation energy $E_{A,g}$, were required to be inputted in gPROMS, and hence, 1 and 3 were attained from previous work done (Ciardha et al., 2012) on paracetamol-ethanol systems, while 2 was approximated to be of order 1 to maintain linearity and simplicity of the kinetic model.

Table 1 summarises the key parameters specified for

Table 1. Parameters specified for the energy balance, nucleation and growth models in the MSMPR crystalliser in gPROMS®

	gPROMS® model parameters	
	Parameter	Value
Energy Balance	Overall Heat Transfer Coefficient	200 $Wm^{-2}K^{-1}$
	Heat Transfer Area	0.01131 m^2
Secondary Nucleation (Attrition)	k_n (CI)	15 $\ln(\#/m^3s)$
	L_{min} (CI)	100 μm
	n (CI)	1
	k_n (CC)	30 $\ln(\#/m^3s)$
	L_{min} (CC)	100 μm
	n (CC)	1
Growth and Dissolution	k_g	10 m/s
	g	1.33
	$E_{A,g}$	45 kJ/mol

the energy balance, Evan's model and the growth and dissolution model in gPROMS®.

4. RESULTS AND DISCUSSION

4.1. Induction time

The induction time obtained from the homogeneous crystallisation experiments was 32 minutes from reaching 20°C, on average. As such the seeds were introduced to the system at 20 minutes, accounting for the observed deviation in induction time to ensure no primary nucleation occured. This step was crucial for the optimisation of seed timing, which ultimately effects the overall process.

4.2. Experiment vs Simulation

The results from the experimental batch crystallisation and the corresponding gPROMS® simulation can be seen in figure 5. It can be observed that the results obtained for milled seeded operations via experiment and modelling gave a larger median CSD compared to the un-milled seeded runs.

However, a notable discrepancy was detected between the two results - the gPROMS® model generated a bimodal particle size distribution of crystals for both seed types unlike the experimental CSDs which were unimodal. It was postulated that this could be due to the limitation of the growth and dissolution model that registers the fragments formed via attrition as initial seeds introduced in the system but of a smaller size. As the growth rate of fragments decrease with decreasing size, smaller fragments are modelled to have very small or almost negligible growth rates (Gahn and Mersmann, 1999b). Thus leading to the accumulation of particles in the region $\leq 50\mu m$

4.3. Milled vs un-milled

Another subtle yet important observation that can be made is that the simulation for un-milled seeds gave a larger median PSD compared to the experimental PSD. This follows on from the limitation mentioned in section 3.2.2 where gPROMS® assumes that seeds have been introduced to the system at t=0 and implies that seeds have a longer residence time in the crystalliser than actually is the case, allowing more time for them to grow and dissolve. It is rather surprising that the same result is not seen with the milled seeds even though all the physical conditions entered were identical in both the simulations and the aforementioned limitation still holds. The only difference was the way the crystal geometry was specified for the milled seeds – equivalent to a quarter of a sphere.

A point to note here is that the median size of the unmilled experimental product was lower than the minimum size of the initial seeds and this was speculated to have arose from the activated secondary nucleation on the seeds. It can be expected that the effect of secondary nucleation on unmilled seeds is higher than in the case of milled ones due to the slower growing facets (Clontz et al., 1972). This lower median size implies that despite seeding under metastable conditions there is still considerable competition between growth and secondary nucleation when unmilled seeds are used. This further highlights the benefits of milled seeds for producing large crystals.

From these observations the inconsistencies in the results can be narrowed down to 2 main reasons; firstly, the limitation of the growth and dissolution model and secondly, the crystal geometry.

5. SELF-HEALING

The gCRYSTAL® simulation was parameterised to replicate the experimental conditions as closely as possible. The inability to specify surface chemistry of seeds in the simulation is an acknowledged limitation of this software, as such, the growth kinetics was investigated. gCRYSTAL's® kinetic model for growth uses the popular two-step classical process, where a the solute must first diffuse from the bulk to the interface, after which it integrates into the crystal lattice. Considering alternative mechanisms for growth; the earliest attempt at describing the free surface energy dependent growth rate of crystals was made by Gibbs(1878). He likened the growth of a crystals to a special case of droplet formation – postulating that a crystal in a supersaturated solution at constant temperature and pressure would grow to minimise the total free surface energy

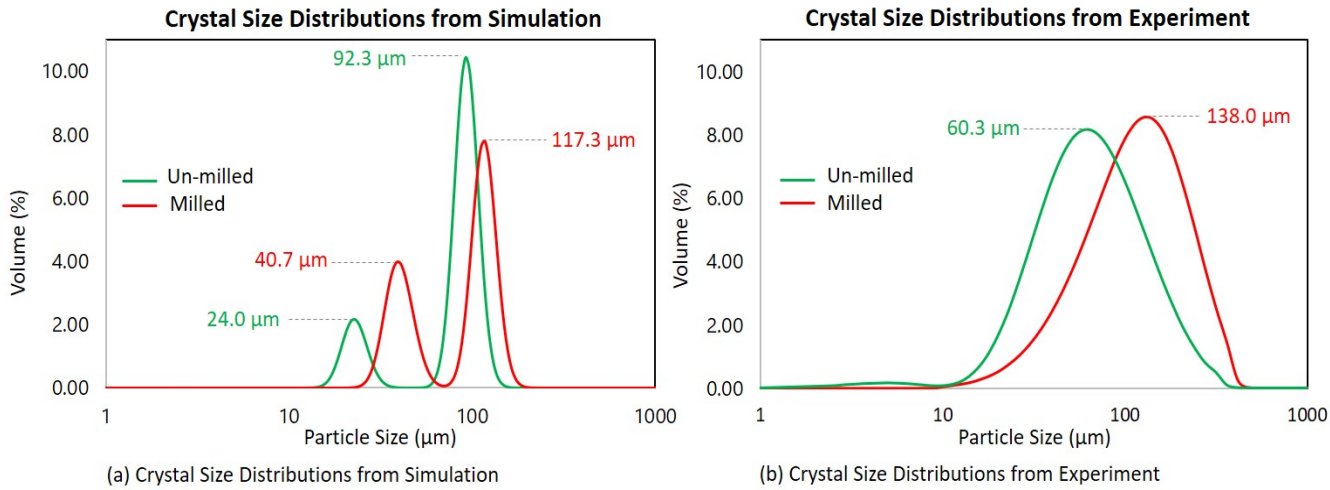


Fig. 5. Particle size distribution for milled and un-milled seeds from (a) Simulated and (b) experimental batch crystallisation

for a given volume(Mullin, 2001). This was further developed by Wulff(1901) with his theorem, resulting in the Wulff construction for determining the equilibrium shape of crystals:

'The minimum surface energy for a given volume of a polyhedron will be achieved if the distances of its faces from one given point are proportional to their capillary constants'.

The term capillary constant here refers to the surface tension. This inspired a thought experiment from first principles; about how two crystals of the same size (one spherical, the other hemispherical and cut along its cleavage plane) would grow in a supersaturated environment. This led to the proposition of a "self-healing" mechanism to explain the observed discrepancy, in accordance with Wulff's finding; that the hemispherical crystal would effectively regrow from the cleavage plane to a sphere before further growth, and it would do so quicker than the crystal which had not been cut.

5.1. Single crystal growth

The self-healing hypothesis was investigated using a paracetamol and ethanol system similar to the case of the batch crystallisation. Facet characterisation was done using Miller indices (Heng et al., 2006).

5.1.1. Crystal Preparation

The paracetamol crystals used for the investigation were obtained using the same method as the milled seeds in 3.1.3, however, they were not crushed.

5.1.2. Crystal Setup

A paracetamol crystal was cut in half along its cleavage plane facet (001), exposing the high energy facets (010). A 35 ml vile was filled up about 80% wih a saturated solution of paracetamol in ethanol and covered with

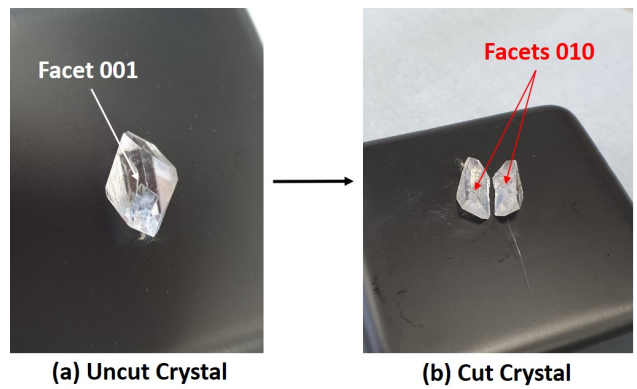


Fig. 6. Facets on an (a) uncut crystal and (b) a crystal cut from its cleavage plane

paraffin wax. An incision was made on the paraffin wax and one half of the crystal was then suspended by kevlar fibres in the vile. It was then covered with a cap (not flush) to limit evaporation. The vile was then enclosed using cardboard to limit the amount of light reflection in the image; making it easier for the resultant images to be analysed using computer vision code.

5.1.3. Camera Set-Up

Two Nikon D90 cameras were used to capture the profile and side view of the exposed high energy facet 101. They were positioned along an arc constructed on the surface from the center of the vile to ensure the pictures were taken from the same distance as this was a requirement of the code later generated. The magnification on the cameras was set to 150x and additional auto extension tubes of 12mm, 20mm, and 36mm were used to increase the zoom. A timer was then connected to the camera to capture images of the vile every 30 seconds.

The vile was allowed in this position for 24 hours.

5.1.4. OpenCV for crystal growth tracking

The OpenCV library for Python™ programming language is the computer vision and image processing program used to track and trace the perimeter of the crystal, from the images obtained.It was developed *de novo* specifically for this project using a threshold and contour algorithm to isolate and enclose the crystal. After which, a button of known size, captured in a similar set up as the crystal from 5.1.2 was used to calibrate the contour estimation to millimetres.

A limitation to this technique of estimating crystal growth is that it has the potential to yield smaller perimeters on successive frames; this is due to the chips on the crystals surface growing, resulting in a smaller perimeter despite growth. In other rare cases light reflections captured by the camera distort the boundary of the crystal, leading to inaccurate images. However, it is a sufficient analytical tool for this experiment.

5.2. Evaluation with existing models

As facet growth rate is measured perpendicularly to the facet in question(Mullin, 2001); it is clear to see from the snapshots in figure 7 that the growth rate of high surface energy facet 010 is considerably faster than that captured by the front view of the cleavage plane. This is in accordance with crystals growing to minimize free surface energy. The crystal also grows to attain the original shape from which it was cut, further consolidating the self-healing hypothesis.

Mersmann and Gahn presented a quantitative model to explicitly account for brittle fracture, and the subsequent growth and dissolution of fragments in crystallisation processes due crystal-impeller interactions; accounting for the surface energies as an important parameter which corrects the inaccuracies introduced by the description of the attrition process (Gahn and Mersmann, 1999b). Firstly, the model uses a number of assumptions with regard to the loading conditions, material property and elastic stress field in the solid to quantify the process of particle removal (fragmentation) as a function of loading conditions and the mechanical properties of the solid; yielding values for the minimum, maximum, total number of fragments generated as well as the distribution:

$$L_{min} = \frac{32}{2}\frac{\mu^3\Gamma}{H^2K_r} \tag{5}$$

$$L_{max} = \frac{1}{2}\left(\frac{H^{\frac{2}{3}}K_r}{\mu\Gamma}\right)^{\frac{1}{3}} W_p^{\frac{4}{9}} \tag{6}$$

$$N = \frac{\pi}{21}\frac{H^{\frac{1}{2}}K_r^{\frac{3}{4}}}{\alpha\mu^{\frac{3}{4}}\Gamma^{\frac{3}{4}}}\left(\frac{1}{L_{min}^{2.25}} - \frac{1}{L_{max}^{2.25}}\right) W_p \tag{7}$$

The number density of the fragments is then obtained by:

$$q_o(L) = \frac{225}{L_{min}^{-2.25} - L_{max}^{-2.25}} L^{-3.25} \tag{8}$$

Where μ is the shear modulus; H is the hardness of the crystal, K_r is the efficiency factor in the radial direction; Γ corresponds to the fracture surface energy, W_p is the elastic plastic energy, and α, the volume shape factor. Emprirical data is required to derive the values of Γ and K_r; these have also been correlated and have been shown to be proportional to the product of the Youngs's modulus and atomic distance:

$$\frac{\Gamma}{K_r} \propto El_o \tag{9}$$

The equation for the growth kinetics of the fragments uses the a similar two step mechanism as mentioned previously, accounting for the integration step as second order with respect to the supersaturation; and can be derived (Mersmann et al., 1992):

$$\frac{G(L)}{2k_d(L)} = \frac{\Delta c}{c_s} + \frac{k_d(L)}{2k_rc_s} - \sqrt{(\frac{k_d(L)}{2k_rc_s})^2 + (\frac{k_d(L)}{k_rc_s})\frac{\Delta c}{c_s}} \tag{10}$$

The driving force for the crystal growth is the degree of supersaturation:

$$\Delta c = c - c* \tag{11}$$

Although the saturation concentration $c*$ is usually taken to be constant; in order to account for the anomalous growth of fragments, the real saturation concentration is equated to the molar strain energy by:

$$c*_{real} \sim c * \exp(\frac{w_s}{RT}) \tag{12}$$

Finally, applying Rittinger's Law, which states that - "the strain energy experienced by a particle during its formation is inversely proportional to its size and not a function of the intensity of impact":

$$\bar{w}L = 2\frac{\Gamma}{K_r} \tag{13}$$

with an assumption of proportionality between the energy stored in the deformed region and inner core of crystal to the ratio of fracture surface and volume of fragment (Gahn and Mersmann, 1999a); the following equation could be obtained, relating the surface energy increase due to attrition of fragments to the 'real' solubility of such fragments:

$$c*_{real} = c * \exp(\frac{\Gamma_s}{RTL}) \tag{14}$$

Thus the key role of Γ_s is to divide the population of fragments into two classes - the large mother crystal

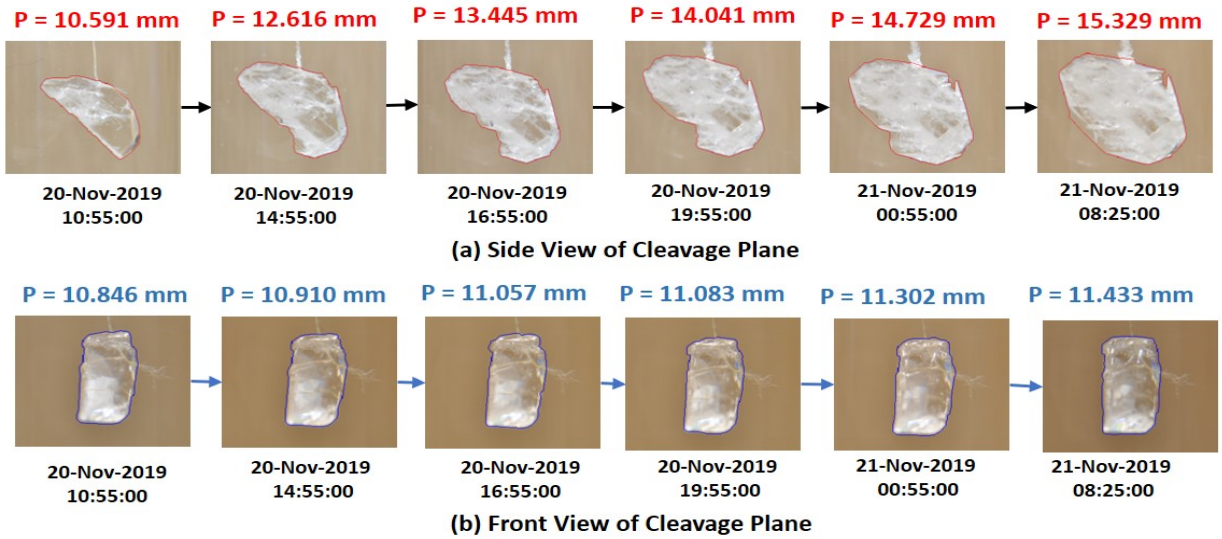


Fig. 7. Snapshots of the (a) Side view and (b) Front View of the cleavage plane of crystal over a 24 Hour Period

and the small growing or dissolvng fragments, explicitly accounting for the growth of the latter. As opposed to the conventional assumption of negligible growth rates of small fragments, which results in the accumulation of fragments in the region $< 50\mu m$ as observed in Fig 5a

6. CONCLUSION AND DISCUSSION

The primary aim of this paper was two-fold; first to engineer large uni modal crystals in a seeded batch reactor, and secondly, to investigate the agreement of the resultant distribution with industry modelling software gCRYSTAL®. The experimental runs which were simulated used seeds of the same size but obtained from two different sources -milled and un-milled. It has been shown by both experimentally obtained CSDs as well as those obtained via simulation that milled seeds, in general, produce larger median size particles. This was attributed to the faster growth kinetics due to the surface chemistry of broken crystals as a result of mechanical stress. These higher energy facets were also shown to reduce the prevalence of secondary nucleation further resulting in larger growth.

The self-healing hypothesis subsequently investigated for a single crystal growth proved analogous to the results of the batch experiments; with milled seeds growing faster and consequently larger than corresponding un-milled ones, for the given residence time. This emphasises that along with growth arising from the adsorption layer theories described in section 2.2 there is also a considerable effect of surface energies on growth which should be considered in conjunction with the former in growth models.

Although the effect of facet surface energies was investigated on a macro-scale and for a longer residence time; there is no indication that the fractured surfaces of large and small fragments should differ as shown by Mersmann and Gahn. In fact, the effect of surface related energy increase with regard to the dissolution and growth is greater on a smaller scale, as is indicative by Rittinger's law. Hence, this effect in the batch crystalliser would actually be more considerable than was observed in the self-healing investigation; provided the fragments are not so small as to result in dissolution.

7. OUTLOOK

From the above findings, the necessity to incorporate quantitative models that account for the effect of surface energies on the growth of fragments, into existing modelling software is evident. Bosetti and Mazzoti recently developed population balance models incorporating the Gahn and Mersmann mechanistic description of attrition in a breakage and nucleation model, separately(Bosetti and Mazzotti, 2019). These models were shown to both be accurate at determining the resultant crystal properties, however, they are dependent on the prevailing mechanism of crystallisation - typically the level of supersaturation and stirring intensity. The effect of surface energies could also be investigated on a smaller scale, using a micro reactor to ascertain the size-independence alluded to.

Analytical techniques developed for this project are versatile; providing the option of refinement for application in both similar and other instances where precise contour tracking for modelling is required.

NOMENCLATURE

α	Volume shape factor, [-]
Δc	Concentration difference, [-]

ϵ	Energy dissipation rate, $[m^2/s^3]$
Γ	Fracture surface energy, $[J/m^2]$
Γ_s	Surface related energy increase, [J m/mol]
μ	Shear modulus, $[N/m^2]$
$\overline{w}$	Average volumetric strain energy, $[J/m^3]$
ρ_c	Crystal density, $[kg/m^3]$
ρ_{crys}	Molar density of the crystal phase, $[mol/m^3]$
σ	Relative supersaturation, $(= \frac{C-C^*}{C^*} = S-1)$, [-]
$\ln k_{n_cc}$	Natural logarithm of secondary nucleation rate constant for crystal-crystal collisions, $[\ln(\#/m^3 s)]$
$\ln k_{n_ci}$	Natural logarithm of secondary nucleation rate constant for crystal-impeller collisions, $[\ln(\#/m^3 s)]$
c	Concentration, $[kmol/m^3]$
c^*	Equilibrium concentration, $[kmol/m^3]$
c_s	Molar concentration, $c_s = \tilde{n}_s$, $[kmol/m^3]$
C_{bulk}	concentration of solute in the bulk of the liquid phase, $[mol/m^3]$
C_{int}	Concentration of solute in the liquid phase at the crystal surface, $[mol/m^3]$
C_{sat}	Saturation concentration of solute in the liquid phase, $[mol/m^3]$
E	Young's modulus, $[N/m^2]$
$E_{A,g}$	activation energy for surface integration, [J/mol]
G	Linear growth rate of crystals, [m/s]
g	supersaturation dependency of the surface integration rate, [-]
H	Hardness, $[N/m^2]$
J_{sec_cc}	Secondary nucleation rate due to crystal-crystal collisions, $[\#/m^3 s]$
J_{sec_ci}	Secondary nucleation rate due to crystal-impeller collisions, $[\#/m^3\, s]$
k_d	Mass transfer coefficient, [m/s]
k_g	proportional factor for surface integration, [m/s]
K_r	Radial efficiency, [-]
k_r	Integration rate constant, $[m^4/(kmol s)]$
k_v	Volume shape factor, [-]
L	Mean size of particles, [m]
l_0	Atomic distance, [m]
L_{max}	Maximum size of the fragments generated, [m]
L_{min_cc}	particle size above which crystals are prone to attrition for crystal-crystal collisions, [m]
L_{min_ci}	particle size above which crystals are prone to attrition for crystal-impeller collisions, [m]
L_{min}	Minimum size of the fragments generated, [m]
N	Number of particles, [-]
n	Number density of particles, $[\#/m^3 m]$
N_p	Impeller power number, [-]
N_Q	Impeller flow number (primary), [-]
n_{cc}	supersaturation dependency of the secondary nucleation rate for crystal-crystal collisions, [-]
n_{ci}	supersaturation dependency of the secondary nucleation rate for crystal-impeller collisions, [-]
q_0	Number density, [1/m]
R	GAs constant, [J/mol K]
T	Absolute temperature, [K]
W	Energy, [J]
W_p	Impact energy, [J]
w_s	Molar strain Energy, [J/mol]

REFERENCES

Agrawal, S. G. and A. H.J. Paterson (2015). "Secondary Nucleation: Mechanisms and Models". In: *Chemical Engineering Communications* 202.5, pp. 698–706. ISSN: 15635201.

Artusio, Fiora and Roberto Pisano (2018). "Surface-induced crystallization of pharmaceuticals and biopharmaceuticals : A review". In: *International Journal of Pharmaceutics* 547.1-2, pp. 190–208. ISSN: 0378-5173.

Bosetti, Luca and Marco Mazzotti (2019). "Population Balance Modeling of Growth and Secondary Nucleation by Attrition and Ripening". In:

BrandConnect (2010). *MSMPR Crystallization Equipment*.

Ciardha, C T O et al. (2012). "Determination of the Crystal Growth Rate of Paracetamol As a Function of Solvent Composition". In:

Clontz, Norvin A. et al. (1972). "Growth of Magnesium Sulfate Heptahydrate Crystals from Solution". In: *Industrial and Engineering Chemistry Fundamentals* 11.3, pp. 368–373. ISSN: 01964313.

CRODA (n.d.). *Lazer Differaction Particle Sizing (Malvern Mastersizer)*.

Evans, T W and A F Sarofim (1974). "Models of Secondary Nucleation Attributable". In: 20.5, pp. 959–966.

Frawley, Patrick J. et al. (2012). "The effects of supersaturation, temperature, agitation and seed surface area on the secondary nucleation of paracetamol in ethanol solutions". In: *Chemical Engineering Science* 75, pp. 183–197. ISSN: 00092509.

Gahn, C. and A. Mersmann (1999a). "Brittle fracture in crystallization processes Part A. Attrition and abrasion of brittle solids". In: *Chemical Engineering Science* 54.9, pp. 1273–1282. ISSN: 00092509.

— (1999b). "Brittle fracture in crystallization processes Part B. Growth of fragments and scale-up of suspension cyrstallizers". In: *Chemical Engineering Science* 54.9, pp. 1283–1292. ISSN: 00092509.

Gao, Zhenguo et al. (2017). "Recent Developments in the Crystallization Process: Toward the Pharmaceutical Industry". In: *Engineering* 3.3, pp. 343–353. ISSN: 20958099.

Granberg, Roger A. and Åke C. Rasmuson (1999). "Solubility of paracetamol in pure solvents". In: *Journal of Chemical and Engineering Data* 44.6, pp. 1391–1395. ISSN: 00219568.

Heng, Jerry Y Y et al. (2006). "Anisotropic Surface Energetics and Wettability of Macroscopic Form I Paracetamol Crystals". In: 21, pp. 2760–2769.

Khabriev, I. Sh et al. (2017). "Some Thermodynamic Characteristics of Paracetamol Dispersing with the SEDS Method". In: *Russian Journal of Physical Chemistry B* 11.8, pp. 1231–1245. ISSN: 19907923.

Kobari, M., N. Kubota, and I. Hirasawa (2010). "Simulation of metastable zone width and induction time for a seeded aqueous solution of potassium sulfate". In: *Journal of Crystal Growth* 312.19, pp. 2734–2739. ISSN: 00220248.

Kubota, Noriaki et al. (2001). "Seeding policy in batch cooling crystallization". In: pp. 31–38.

Li, Huayu et al. (2017). "Modeling of Nucleation and Growth Kinetics for Unseeded Batch Cooling Crystallization". In: *Industrial and Engineering Chemistry Research* 56.14, pp. 4060–4073. ISSN: 15205045.

Li, L. and N. Rodríguez-Hornedo (1992). "Growth kinetics and mechanism of glycine crystals". In: *Journal of Crystal Growth* 121.1-2, pp. 33–38. ISSN: 00220248.

Mullin, John W (2001). *Crystallization*. Ed. by Elsevier. Fourth Edi.

Myerson, Allan S. (2001). *Handbook of Industrial Crystallization*. 2nd, p. 304.

Nguyen, T T H et al. (2014). "Precision measurement of the growth rate and mechanism of ibuprofen {001} and {011} as a function of crystallization environment †". In: pp. 4568–4586.

Parsons, A. R., S. N. Black, and R. Colling (2003). "Automated measurement of metastable zones for pharmaceutical compounds". In: *Chemical Engineering Research and Design* 81.6, pp. 700–704. ISSN: 02638762.

PSE Model Documentation (2015). "gCRYSTAL Crystalliser MSMPR Fixed Volume". In: pp. 1–46.

Simone, Elena (2018). "Preventing Crystal Agglomeration of Pharmaceutical Crystals Using Temperature Cycling and a Novel Membrane Crystallization Procedure for Seed Crystal Generation". In:

Su, Qinglin, Zoltan K Nagy, and Chris D Rielly (2015). "Chemical Engineering and Processing : Process Intensi fi cation Pharmaceutical crystallisation processes from batch to continuous operation using MSMPR stages : Modelling , design , and control". In: 89, pp. 41–53.

UK Government (2017). "Made Smarter Review". In: p. 231.

Verdurand, Elodie et al. (2005). "Secondary nucleation and growth of organic crystals in industrial crystallization". In: *Journal of Crystal Growth* 275.1-2, pp. 1363–1367. ISSN: 00220248.

Virone, C. et al. (2005). "Growth rate dispersion of ammonium sulphate attrition fragments". In: *Journal of Crystal Growth* 275.1-2, pp. 1397–1401. ISSN: 00220248.

Enviro-economic analysis of the production of Dimethyl Ether via Syngas

Maxim S.F. Bird & Bohan Xu
Supervisor: Prof. Benoit Chachuat
Department of Chemical Engineering, Imperial College London

Abstract

Dimethyl Ether (DME) has shown potential to be a more sustainable alternative fuel to LPG and diesel. However, in industry it is generally synthesised from syngas in an unsustainable process. Therefore, it is of particular interest to analyse the various different production routes to establish an economically-feasible, environmentally-friendly pathway. Combinations of dry methane reforming (DMR), steam methane reforming (SMR), autothermal reforming (ATR) and reverse water gas shift (RWGS) processes as well as biomass gasification are explored for the synthesis of syngas with ratios of 1:1, 3:2 and 2:1. The various syngas products are then reacted both directly and indirectly (via methanol) to form DME. The processes are modelled in Aspen HYSYS and a life cycle assessment (LCA) is carried out using the program Simapro. An enviro-economic comparison of these process alternatives is presented in this work. The indirect production routes of DME were established to be the more sustainable. Of these the most profitable used syngas produced in a SMR with RWGS process. The effects of DME price and enterprise discount rate on the economic performance and an analysis of the causes of the environmental impact of this process are also demonstrated.

Keywords: [Dimethyl Ether (DME), Syngas, Aspen HYSYS, Life Cycle Assessment (LCA), Simapro]

1 Introduction

1.1 Motivation

Dimethyl ether (DME) is primarily used in industry as a low-temperature solvent and extracting agent (International DME Association n.d.). However, it has shown major promise as an alternative fuel to both LPG in the HI&I industry and diesel in specifically-designed, compression ignition engines (Alternative Fuels Data Center n.d.). This is reinforced by the fact that there have been a number of DME vehicle demonstrations including one in which ten vehicles were operated for 750,000 miles (Alternative Fuels Data Center n.d.). Furthermore, it is easily liquefied and transported. Due to its short carbon chain structure, far fewer particulate emissions are released upon its combustion. Additionally, according to LCA analysis, it is approximately five times more energy efficient than traditional fuels making it an environmentally friendly substitute. DME's other uses include as an aerosol, a propellant or a refrigerant (Pubchem n.d.).

In industry DME is generally synthesised from syngas - a mixture of carbon monoxide (CO) and hydrogen (H_2), which in-turn is produced from natural gas or the gasification of biomass. Usually, this production process is performed in an unsustainable way. Therefore the aim of this research is to find a environmentally-sustainable, economically-feasible method for DME production from syngas.

1.2 Background

Sustainable routes for the production of DME and syngas using the following methods have been researched previously although these methods were not considered in this research.

Solar thermochemical processes

Syngas can be produced from carbon dioxide (CO_2) and water in a solar powered thermochemical reaction. This sustainable process operates on the principle of a transition between the oxidized (higher-valence, $MeO_{oxidized}$) and reduced (lower-valence, $MO_{reduced}$) form of a metal oxide exhibiting multiple oxidation states. Examples of these metals include Zn, Cd, Sn and Ge (Agrafiotis et al. 2015).

Electrochemical reduction to syngas

CO_2 and water produced from the combustion of fuel can be reduced to CO and H_2 through electrochemical reduction reactions. An intrinsic environment for the electrolysis process must be set up such that both reduction reactions can happen simultaneously to produce syngas (Simelys Hernández 2017).

Direct hydrogenation of CO_2

The CO_2 can be reacted with H_2 directly to form methanol, which can subsequently be dehydrated to DME. However, a reverse water gas shift reaction will occur simultaneously producing syngas. The reactions can be catalysed by a wide range of metal oxides although the best of these is yet to be established (Suwannapichat et al. 2018).

2 Theory

2.1 Syngas production methods

2.1.1 Steam methane reforming (SMR):

Steam methane reforming is the reaction between steam and methane to form H_2 and CO in a 3:1 ratio if the feed is stoichiometric. The process is carried out over a Ni/Al_2O_3 catalyst with the following reactions taking place (Chibane & Djellouli 2011):

$$CH_4 + H_2O \Longleftrightarrow CO + 3H_2 \quad \Delta H^0_{298} = +206\ kJkmol^{-1}$$

$$CO + H_2O \Longleftrightarrow CO_2 + H_2 \quad \Delta H^0_{298} = -41\ kJkmol^{-1}$$

$$CH_4 + 2H_2O \Longleftrightarrow CO_2 + 4H_2 \quad \Delta H^0_{298} = +165\ kJkmol^{-1}$$

The Langmuir-Hinshelwood mechanism is most appropriate to model the reaction kinetics. However, to simulate the reaction in the Aspen HYSYS Plug model, a more simplified Arrhenius rate law was used to describe the SMR process (Luyben 2016):

$$r_1 = 0.1exp(\frac{-275000}{RT})P_{CH_4}P_{H_2O} - 9.7x10^{-14}exp(\frac{-47400}{RT})P_{CO}P_{H_2} \quad (1)$$

$$r_2 = 2.6x10^{-7}exp(\frac{-47400}{RT})P_{CO}P_{H_2O} - 1.575x10^{-5}exp(\frac{-85460}{RT})P_{CO_2}P_{H_2} \quad (2)$$

Where pressures are in Pa, temperatures is in K, the reaction is in the vapor phase and the rate has units of kmol·m^{-3}s. The third reaction was assumed to be a sum of the first and second.

2.1.2 Dry methane reforming (DMR):

DMR is the reforming reaction between methane and CO_2 to produce syngas with a H_2:CO ratio of 1:1. During the course of the process the following reactions take place over a Ni/Al_2O_3 catalyst (Benguerba et al. 2015):

$$CH_4 + CO_2 \Longleftrightarrow 2CO + 2H_2 \quad \Delta H^0_{298} = +247\ kJkmol^{-1}$$

$$CO_2 + H_2 \Longleftrightarrow CO + H_2O \quad \Delta H^0_{298} = +41.7\ kJkmol^{-1}$$

However, since all the reactions are highly endothermic, a large amount of carbon may be produced. Therefore, the reactor is simulated with the following rate laws at a temperature such that the rate of these side reactions will be sufficiently low as to be ignored:

$$r_1 = 2000exp(\frac{-300000}{RT})P_{CO_2}P_{CH_4} - 2.738x10^{-11}exp(\frac{-48513}{RT})P_{CO}P_{H_2} \quad (3)$$

$$r_2 = 10^{-11}exp(\frac{-50000}{RT})P_{CO_2}P_{H_2} \quad (4)$$

Where pressures are in Pa, temperatures in K, and the reaction is in the vapour phase and has rate units of kmol·m^{-3}s.

2.1.3 Autothermal reforming (ATR):

ATR is a process in which a natural gas feed is partially oxidised by an oxygen and steam stream in a Ni/Al_2O_3 catalysed reformer (Air Liquide 2019). The feedstock can also be refinery off gas, pre-reformed gas, Fischer-Tropsch tail-gas or Liquefied Petroleum Gas naphtha. The feed gas is first desulfurised and then pre-heated before entering the ATR reactor at a pressure range between 30 and 100 barg (Air Liquide 2019). The same reforming reactions take place during the ATR process as in the SMR process. However, due to the presence of oxygen,the direct oxidation of methane also occurs (Hoang & Chan 2014):

$$CH_4 + 2O_2 \Longleftrightarrow CO_2 + 2H_2O \quad \Delta H^0_{298} = -802\ kJkmol^{-1}$$

The rate laws used for the process simulation were as follows:

$$r_1 = \frac{k_1 P_{CH_4} P_{O_2}{}^{\frac{1}{2}}}{(1 + K^C{}_{CH_4}P_{CH_4} + K^C{}_{O_2}P_{O_2}{}^{\frac{1}{2}})^2} \quad (5)$$

$$r_2 = \frac{k_2}{P_{H_2}{}^{2.5}}(P_{CH_4}P_{H_2O} - \frac{P_{H_2}{}^3 P_{CO}}{K_{e^2}}) \times \frac{1}{Q_r{}^2} \quad (6)$$

$$r_3 = \frac{k_3}{P_{H_2}}(P_{CO}P_{H_2O} - \frac{P_{H_2}P_{CO_2}}{K_{e^3}}) \times \frac{1}{Q_r{}^2} \quad (7)$$

$$r_4 = \frac{k_4}{P_{H_2}{}^{3.5}}(P_{CH_4}P^2_{H_2O} - \frac{P_{H_4}{}^3 P_{CO_2}}{K_{e^4}}) \times \frac{1}{Q_r{}^2} \quad (8)$$

$$Q_r = 1 + K_{CO}P_{CO} + K_{H_2}P_{H_2} + K_{CH_4}P_{CH_4} + \frac{K_{H_2O}P_{H_2O}}{P_{H_2}} \quad (9)$$

2.1.4 Reverse water gas shift (RWGS):

Reverse water gas shift is a side reaction which can be used in syngas production in order to lower the H_2:CO ratio. The process requires a Fe/Al_2O_3 catalyst and the reaction is the same as the second reaction in the DMR process (Benguerba et al. 2015).

2.1.5 Gasification of Biomass/Coal

The process steps involved in biomass or coal gasification to produce syngas are drying, pyrolytic decomposition, oxidation and reduction (Trippe et al. 2011). During the first step, the slurry is heated up to 200 °C to boil off any present moisture.Then the mixture is heated further to allow the thermal or pyrolytic decomposition to take place. Oxygen and steam are introduced to allow for oxidation to occur. This process is highly exothermic generating temperatures of up to 2000 °C, which provides the heat required for the process. The stream is then cooled to a temperature between 800 °C and 1100 °C for reduction to occur and form the syngas product.

2.2 DME Production Process

According to Shen (2019), DME can either be produced from syngas directly or indirectly (via Methanol). During these processes the following reactions occur (Ng et al. 1999):

$$CO + 2H_2 \Longleftrightarrow CH_3OH$$

$$CO_2 + 3H_2 \Longleftrightarrow CH_3OH + H_2O$$

$$CO + H_2O \Longleftrightarrow CO_2 + H_2$$

$$2CH_3OH \leftrightarrow CH_3OCH_3 + H_2O$$

For the purpose of simulating these processes the following rate laws were assumed (Ng et al. 1999):

$$r_{CO_2 hydrogenation} = \frac{k_1(P_{H_2}P_{CO_2})[1 - \frac{(\frac{1}{K_{eqm1}})(P_{CH_3OH}P_{H_2O})}{P_{CO}P^3_{H_2}}]}{(1 + K_2\frac{P_{H_2O}}{P_{H_2}} + \sqrt{K_3P_{H_2}} + K_4P_{H_2O})^3} \quad (10)$$

$$r_{RWGS} = \frac{k_5P_{CO_2}[1 - \frac{(\frac{1}{K_{eqm2}})(P_{CO}P_{H_2O})}{P_{CO}P_{H_2}}]}{1 + K_2\frac{P_{H_2O}}{P_{H_2}} + \sqrt{K_3P_{H_2}} + K_4P_{H_2O}} \quad (11)$$

$$r_{MeOHdehydration} = k_6 K_{CH_3OH}{}^2 \frac{C_{CH_3OH}{}^2 - \frac{C_{H_2O}C_{DME}}{K_{eqm3}}}{(1+2\sqrt{K_{CH_3OH}}C_{CH_3OH}+K_{H_2O}C_{H_2O})^4} \quad (12)$$

2.2.1 Direct DME production process

The direct process takes place in an isothermal, fluidized-bed, catalytic reactor at 260°C with a feed pressure of 54.9bar and a pressure drop of 10bar. The outlet of the reactor is cooled down to 35°C and the water subsequently flashed out. A sequences of distillation columns is then used to separate the high percentages of methanol and the DME, The methanol stream and the unreacted gaseous stream are fed back to the reactor to increase the overall conversion of the process.

2.2.2 Indirect DME production process

Conversely during indirect DME production, there are two reactors. Methanol is synthesised in the first isothermal fluidized bed catalytic reactor at 260°C. A distillation column then separates a high purity methanol stream from a syngas stream, which is recycled. The methanol is subsequently fed to a second fluidized bed catalytic reactor at 0.6bar and 290°C. The reactor outlet temperature is controlled to be 360°C, to promote a healthy rate of reaction. The products then pass into a sequence of distillation columns first producing a high purity DME stream, then separating the remaining methanol from the residual water to be recycled.

3 Methodology

Various reaction processes were simulated to produce syngas with H_2:CO ratios of 1:1, 3:2 and 2:1. The economic and environmental performances of these processes along with others sourced from Trippe et al. (2011) and Baltrusaitis & Luyben (2015) were then compared. The processes showing the best results for each of the syngas ratios were then modelled as the feedstock in the direct (and in the case of 2:1 syngas the indirect) production of DME. The economic and environmental performance comparison were then repeated for the DME production. In this way a preferable production route was established.

3.1 Process Simulations

Aspen HYSYS was used to simulate the different syngas production routes, calculating the waste flow rates and compositions, energy consumptions, and the capital and operating costs for each process.

3.1.1 DMR only

To produce a syngas stream with a 1:1 H_2:CO ratio, the DMR route is chosen according to the stoichiometric ratio of the main reactants. During this process, the CO_2 is compressed to 4.3bar and mixed with methane at the same pressure. The mixed gas is heated to 900°C in a fired heater before sending to the fixed bed catalytic reactor at 1000°C. After that the outlet stream is cooled to 50°C by two coolers in series, the first generating high-pressure steam. Then the outlet stream is expanded to 2.5bar and cooled down to 25°C in order to meet the requirement of feed to the DME production plant.

3.1.2 ATR only

During the ATR process, the water is pumped up to 34bar and mixed with methane at the same pressure. The mixed stream is then heated to 200°C before feeding into a fluidized bed reactor. The oxygen stream at 200°C and 34bar is fed into the reactor at the same time. The reactor is heated up to 1850°C by fire heater and the pressure is dropped to 12bar for the reaction to have an appropriate rate and selectivity. The outlet of the reactor is cooled down to 10°C by three heat exchangers generating high and low pressure steam and reduce the demanding for refrigerants. The water is then flashed out and the syngas stream is compressed and re-heated to reach the conditions required for DME production. By tuning the ratio of water, methane and oxygen, syngas ratios of H_2 and CO at 2:1 and 3:2 can be achieved.

3.1.3 SMR combined with DMR

Since SMR produces syngas with a H_2:CO ratio of 3:1 or higher, whilst DMR produces it in a ratio of 1:1, a combination of the two processes in parallel can result in a stream with a 2:1 or 3:2 ratio. During the SMR process, the water is pumped up to 34bar and mixed with methane at the same pressure. The mixture is then heated to 900°C before being fed into a fixed bed catalytic reactor. The reactor itself is heated to 1500°C. The outlet stream from the reactor is cooled down to 10°C. Subsequently, the water in the stream is flashed out, before the syngas product is mixed with the product from the DMR process.

3.1.4 SMR combined with ATR

A SMR reactor can likewise be combined with an ATR reactor to lower the H_2:CO ratio. However, these reactors take place in series not parallel. According to the literature (Baltrusaitis & Luyben 2015), the methane is fed into the SMR reactor first with a steam feed in approximately a 10:1 ratio at 400°C. This SMR reactor is heated to 600°C. The outlet is then fed into the ATR reactor with a large oxygen stream. However, when this model was simulated in Aspen HYSYS, the product streams predominantly consisted of CO_2 and water. This was due to the lack of water in the feed stream to the ATR reactor causing oxidation to be promoted rather than reforming. Therefore, the two processes were instead connected in parallel and by adjusting their relative production rates, the desired H_2:CO ratios were produced.

3.1.5 SMR combined with RWGS

An alternative way to lower the SMR H_2:CO ratio is to have a RWGS reactor in series. In this reactor, the H_2 is partially consumed and more CO is produced. Therefore, by tuning the relative inlet methane and CO_2 flowrates, ratios of 2:1 and 3:2 can be achieved.

3.1.6 DME synthesis

The processes modelled for the production of DME from syngas are based on the work previously carried out by Shen (2019). However, minor edits were made to the sizing of the reactors and the energy integration system for both the direct and indirect process models. Additionally, for the direct production route, the process was adjusted to have a syngas feed for each of the different H_2:CO ratios and a Rankine cycle was added to improve the energy efficiency of the process by making use of the reactor heat.

3.2 Energy Integration

The simulations performed in Aspen HYSYS were initially modelled without a heat exchanger netowrk (HEN). Therefore, to reduce the required utility costs and make the designs more competitive, the Aspen Energy Analyzer software was used to find the local HEN optima. This software uses the hot streams that require cooling to heat the cool streams that require heating to design a HEN but takes into account the cost of each utility and the cost of construction of each of the units. Using the resultant capital and operating costs of the designed HEN, the software then performs a basic economic analysis assuming a plant lifetime of 30 years (with an operating time of 8000hrs/yr) to provide a total cost for the design. These local optimum designs were then compared and the design with the lowest total cost selected. The resultant changes in capital and operating costs were then factored into the overall costs.

3.3 Life Cycle Assessment (LCA)

A Life Cycle Assessment (LCA) is a method that quantifies the environmental impacts associated with a product, process or activity. The application of a LCA involves first evaluating the raw materials and energy required for a given process along with the emissions released to the environment. These are then translated into a set of environmental impacts that can be used to compare different design alternatives. The application of a LCA involves the following four phases:

3.3.1 Phase I: Defining the Goal and Scope

During this stage, the system boundaries, impact categories and functional unit for the process are defined. The system boundaries address which upstream and downstream processes are included within the scope of the LCA.

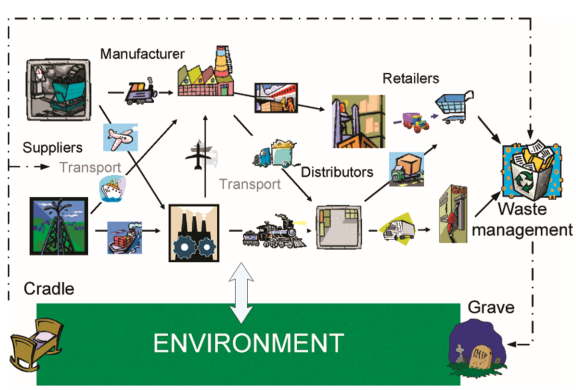


Figure 1: Boundaries of a Life Cycle Assessment (Guillen-Gosalbez et al. 2008)

Since, this paper's aim is to decrease the environmental impact of the production process of DME, the further downstream (e.g. product use and disposal) are neglected. Therefore, the LCA will be a "cradle-to-gate" analysis rather than a full "cradle -to-grave" one. The impact categories are defined by characterization factors which are in turn generally derived at two different levels: the process midpoint (typically the point at which the environmental mechanism is identical for all flows) or the endpoint. At the endpoint level, these characterization factors correspond to three areas: human health, ecosystem quality and resource scarcity. Generally, midpoint analysis has a lower uncertainty and a stronger relation to the environmental flows, whilst endpoint analysis provides better information on the relevance of these flows. In this project endpoint analysis was performed as it facilitates more structured weightings across the damage categories. The functional unit was defined to be per kg of the product of each process to allow for a fair comparison at each stage.

3.3.2 Phase II: Inventory Analysis

The second stage involves collating the required raw materials, energies and waste emissions for the process relative to the size of the previously defined functional unit. Therefore, the required data was taken from the papers by Trippe et al. (2011) and Baltrusaitis & Luyben (2015) and from the simulations designed in Aspen HYSYS. To account for the emissions and energies required for the production of the feedstocks and utilities in the inventory, the Eco-invent v3 APOS database was used.

3.3.3 Phase III: Impact Assessment

The data collated in the inventory is then converted to environmental impacts using the ReCiPe2016 methodology in the Simapro software (Huijbregts et al. 2016). This strategy allows for three perspectives: individualistic, hierarchist and egalitarian using time horizons of 20, 100 and 1000 years respectively. As the hierarchist perspective is based on scientific consensus with regards to future socio-economic developments and its impact plausibility, it was chosen for the LCA analysis. The resultant environmental impacts are split into the aforementioned three categories. Human health damages are measured in disability adjusted life years (DALYs). 1 DALY is equal to a reduction in the lifespan of 1 individual by 1 year or their suffering a disability with a weighting of 0.1 for 10 years. Ecosystem quality is measured in species.year, which is the integrated number of species that will become extinct as a result of the damage caused to their ecosystem over the lifetime of the LCA. Resource scarcity is measured in USD_{2013}. A damage of 1 means that the extraction of a certain resource will have a higher cost by $\$_{2013}1$ in future due to the diminishing reserves.

3.3.4 Phase IV: Interpretation

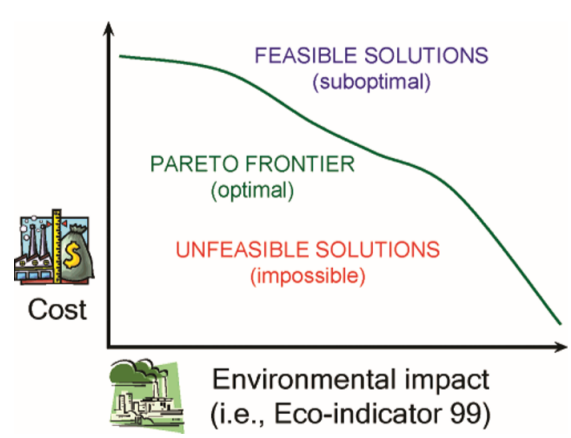


Figure 2: Pareto Curve (Guillen-Gosalbez et al. 2008)

During this stage, the results of the LCA are analysed to compare the different process pathways. This is best performed by the formation of a Pareto curve (see Figure 2) which differentiates between the optimal and suboptimal processes. To plot the Pareto frontier with all the environmental impacts on the same x-axis, each of the different categories needed to be converted into the same units. Therefore, the Human Health and Ecosystem quality categories were converted to monetary values (EUR_{2003}) using the method found in (Weidema 2014). Then taking account for inflation and exchange rates each unit was converted to USD_{2019} to allow for an overall environmental impact (see Table 1).

Impact Category	Unit	$EURO_{2003}$	USD_{2019}
Human Health	DALY	74000	1.170×10^5
Ecosystem Quality	Species.yr	9.500×10^6	1.501×10^7
Resource Scarcity	USD_{2013}	N/A	1.100

Table 1: Factors used to convert environmental impact to USD_{2019}

The production processes that fall on or closest to the Pareto curve are progressed to the next stage of analysis (i.e. only the best performing syngas production processes were used in the DME production analysis)

3.4 Economic analysis

3.4.1 Initial economic analysis

The previously outlined Pareto curve, which was used to compare the production routes for each step of the process, requires an economic calculation to estimate the cost of production. Therefore, since a large number of processes were being compared, a relatively simple model was imposed, the assumptions of which were as follows:

- The plant construction was funded by a loan discounted at a rate of 5%
- The plant would require 18 months to construct
- The loan would be repaid in 10 equal instalments beginning at the end of the first full year of operation
- An inflation rate of 2.5%

The capital costs were estimated in Aspen HYSYS for the simulated models. Meanwhile, the operating costs were estimated by summing the utility costs from Aspen HYSYS with the costs of the required feed and catalysts and assuming 50 operators with a salary of $25000. The feed and catalyst costs were estimated as in Tables 2 & 3 respectively:

	Cost ($/t)	Source
Methane	26920	(Baltrusaitis & Luyben 2015)
CO_2	41.80	(Perez-Fortes & Tzimas 2016)
$O2_2$	80.00	(Baltrusaitis & Luyben 2015)
Steam	0.3963	(American Water Work A n.d.)

Table 2: Feed cost

	Cost ($/kg)	Life (year)	Source
Ni/Al_2O_3	36.88	1	(Millipore Sigma n.d.) (Misc n.d.*c*)
Fe/Al_2O_3	7.394	1	(Misc n.d.*b*)
Cu/ZnO	4.219	4	(Misc n.d.*a*) (Aldibaba n.d.)
α-Al_2O_3	0.7500	5	(Millipore Sigma n.d.)

Table 3: catalyst cost

As for the processes sourced from the literature, the previously calculated costs were used. This economic model resulted in a minimum cost of product that would be required for the plant to break even that was plotted on the Pareto curve. The minimum costs of syngas and methanol that were found to be most promising according to the Pareto curve analysis were then used as the feed costs in the DME production analysis.

3.4.2 Detailed economic analysis

For the most promising DME production processes according to the initial LCA and economic analysis, a more detailed economic analysis was performed to establish and compare the economic performance of the plant. This was necessary as processes that have a high cost of production may still fall on the the Pareto curve of optimal solutions (so long as they have a sufficiently low environmental impact) but may not be able to compete in the current market due to these high costs. The capital investment for each process was estimated using the ISBL method outlined in Towler & Sinnott (2012). Meanwhile, the variable costs for the entire process were established using the same method as previously but a wastewater treatment cost of $0.056/t was also added. The entire process was then modelled using the following assumptions were used in the process:

- A construction time of 3 years financed by a loan with an interest rate of 6%
- At the end of the 3 years the loan is paid off by issuing 10 year bonds at an interest rate of 4%
- Inflation, discount and tax rates of 2.5%, 10% and 27% respectively.
- The TCI is depreciated over the first 10 years of production
- The cost of DME is $1.05/kg
- The plant is operated for 8000hrs/yr

The resultant economic performance of the most promising DME production routes would be compared over a 30 year period and the option with the highest NPV% would be selected. Discounted NPV% was chosen as the parameter for comparison as it gives the best indicator of overall economic performance with respect to the initial capital investment. It can be calculated as follows:

$$NPV\%_n = \frac{NPV_n}{n \times TI} \tag{13}$$

Where n is the number of years since the plant construction began and TI is the total investment.

4 Results and Discussions

4.1 Combined Economic and LCA results

The minimum costs of production and environmental impacts for the various production routes were plotted on Pareto curves for each step of the process to find the most promising process pathway. For the syngas production processes, the results based on the gasification of biomass/coal production method from (Trippe et al. 2011) are plotted in green, likewise those sourced from Baltrusaitis & Luyben (2015) in light blue and the simulated model in black. As can be seen in Figure 4 the ATR, SMR/ATR and SMR/RWGS simulated models had very high environmental impacts. These resulted in the scale of the graphs being increased and the results becoming indistinguishable from one another in Figures 3 and 5. Therefore, these results were not plotted. (They can be found in the supplementary information). These high environmental costs were due to the high requirement for refrigerants in these processes, despite the HEN that had been installed.

4.1.1 1:1 Syngas

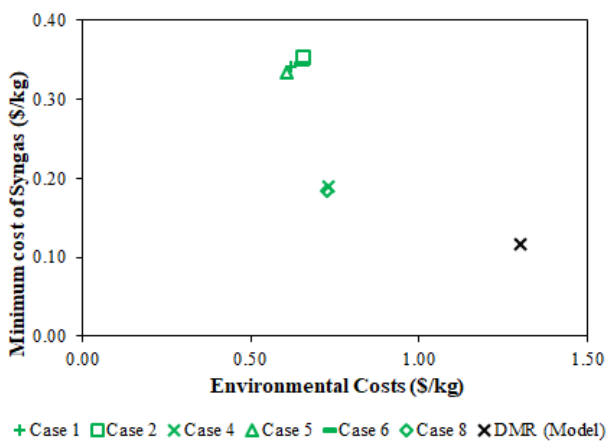


Figure 3: 1:1 Syngas production Pareto curve

The 1:1 syngas production process show the signs of forming a Pareto curve (see Figure 3). (It must be noted that of the simulated models for the production of 1:1 syngas, only the DMR process had a low enough environmental impact to be included). As for the processes designed in Trippe et al. (2011), Case 4 and Case 8 have a much lower cost of production but a higher environmental impact than the others. This is due to the feedstock slurry predominantly being made up of coal rather than biomass, a much cheaper but less environmentally friendly alternative. This change of feedstock resulted in the cost of syngas being almost halved (approximately $0.15 lower) whilst the environmental impact was only increased by a small fraction (approximately $0.10). Therefore, it may be concluded that only the coal gasification processes and DMR model are on the Pareto curve. As Case 8 has a lower cost of production than Case 4 it shows more potential. Therefore, Case 8 and the DMR model production processes were taken forward to the DME production analysis stage.

4.1.2 3:2 Syngas

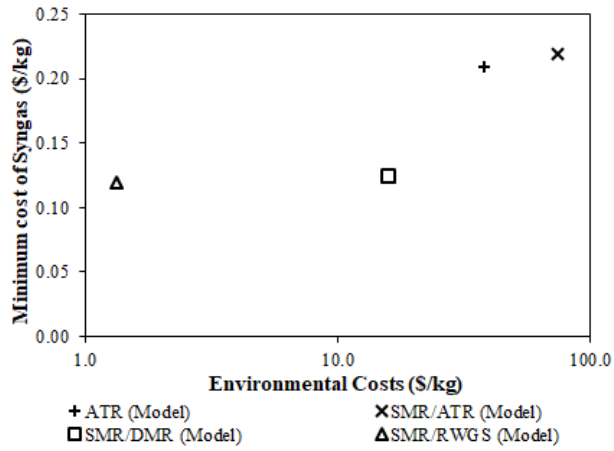


Figure 4: 3:2 Syngas production Pareto curve

In Figure 4 it is evident that no clear Pareto frontier for the production of 3:2 syngas is formed. The ATR reactors modelled were much larger than for the other simulated processes to promote the reforming reaction rather than oxidation. This resulted in a much higher capital expenditure, explaining their high minimum costs of production.The SMR/RWGS process, unlike the others, did not require the use of refrigerants in its production, resulting in its much lower environmental impact. Therefore this process was selected to progress to be analysed for the production of DME.

4.1.3 2:1 Syngas

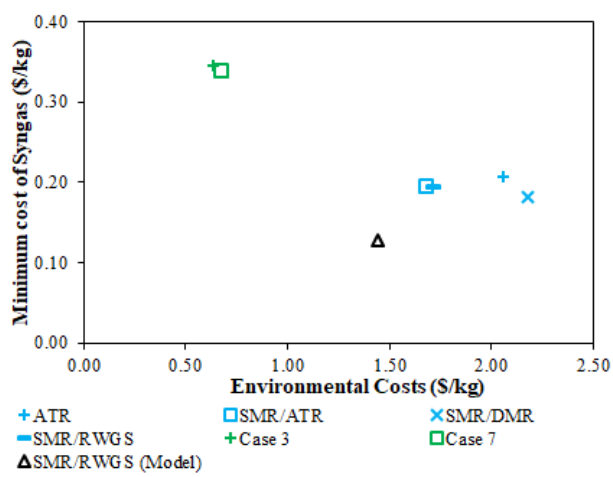


Figure 5: 2:1 Syngas production Pareto curve

Figure 5 shows that the source of the results dictated the relative performance of the process. The biomass gasification processes had the highest cost of production but the lowest environmental impact, due to the assumed straw feedstock being much more sustainable than natural gas. Meanwhile, the simulated models had the lowest cost and generally the highest environmental impact. (N.B. The other

simulated models are not shown due to their very high environmental impact). Once again no clear optimal frontier is produced. However, the biomass production routes and the SMR/RWGS simulation are clearly the most promising results and so were progressed to the Methanol and DME production analyses.

4.1.4 Methanol

As the model for the production of methanol uses 2:1 syngas as the feedstock, the results for the methanol production appeared to be a simple translation of those in Figure 5. Therefore, the same two processes were progressed to the DME production analysis.

4.1.5 DME

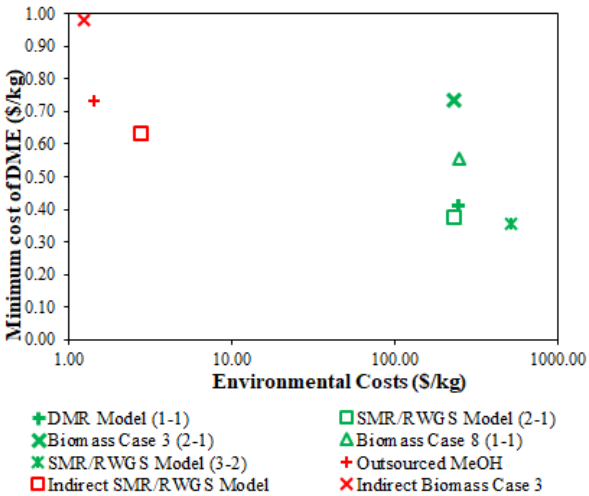


Figure 6: DME production Pareto curve

Figure 6 shows a clear pattern of the indirect production routes (plotted in red) having a higher cost of production but a much lower environmental impact than the direct routes of production (plotted in green). An outline of a Pareto frontier can be seen. However, the benefit of the lower cost of production of the direct processes is far outweighed by their high environmental impact. This is caused by the large number of refrigerants required in the complex separation route in the direct processes. Therefore only the indirect processes were considered for further more detailed economic analysis. However, the production of DME directly from outsourced methanol was not considered for further analysis as it was deemed more inaccurate due to the higher number of assumptions made for this process. These included the price of methanol and the fact that for the LCA performed in Simapro, the methanol input was taken from the Eco-invent database. However, the overall economic feasibility of the indirect processes are called in to question when the price per unit energies are considered. For the minimum cost of production of methanol was $14.55/MJ and $23.36/MJ from syngas synthesised via a SMR with RWGS process and the gasification of biomass, respectively. By contrast, the respective minimum costs of DME production were $19.85/kg and $30.88/kg, representing an increase in cost per unit energy. Therefore, if the DME is to be sold for use as a fuel, an evaluation into the benefits of selling it rather than methanol is necessary.

4.2 Detailed economic comparison

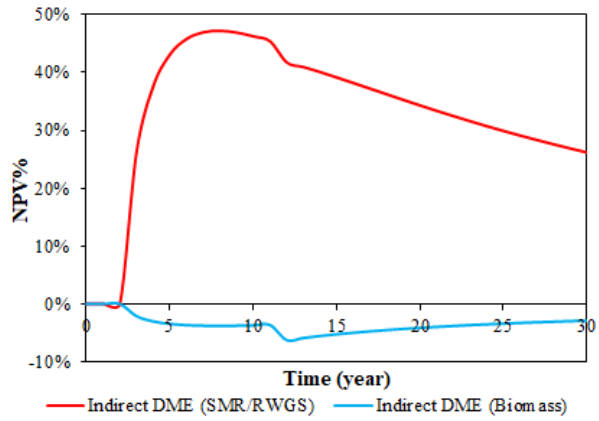


Figure 7: Economic performance of the two most promising DME production routes

According to Figure 7, it is clear that the economic performance of the process producing DME indirectly from syngas produced by the gasification of biomass is much poorer than that using syngas produced in a SMR with RWGS process. Moreover, as the former has a negative NPV% for the entire plant lifetime, it can be concluded that this process may not be competitive in the current market. This is due to the fact that the assumed price of DME of $1.05/kg is insufficient to generate enough revenue to be able to pay off the construction costs of the plant and its operating costs. This is due to the high costs associated with producing the biomass feedstock as it requires a large amount of processing to prepare it into the slurry form appropriate for the gasification process. Since, the production route using syngas produced in a SMR with RWGS process by contrast generates a healthy profit, this production route was preferable. The shape of the economic performance graph for this process can be explained by the assumptions made to perform the economic analysis. The initial steep gradient is due to the fact that discounted NPV% is being calculated, so the further forward in time that the model progresses the more the NPV% is discounted, hence a decreasing gradient. The cause of the sharp drop at year 13 is the paying back of the 10 year bonds. The continuous negative gradient from year 13 onwards is caused by the end of the 10 year period during which the capital investment in the plant was being depreciated resulting in less tax being paid. Therefore, after this period the plant profits are proportionally decreased, corresponding to a lower NPV%, hence the results begin to tend to this lower value.

4.3 Sensitivity Analysis

As the final production route was selected based on its economic feasibility, a sensitivity analysis was required to establish whether the process remained profitable under fluctuating market conditions. Several variables were adjusted

to measure their impact but the price of DME and the enterprise discount rate were found to have the largest impact on the NPV% (the measure of economic performance used).

4.3.1 Price of DME

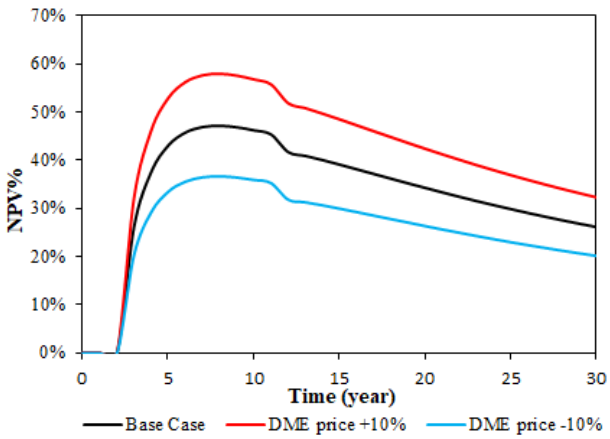


Figure 8: Process economic sensitivity to DME price

The economic performance of the indirect DME production process from syngas produced in a SMR with RWGS process was found to vary quite significantly with fluctuations in DME price as represented in Figure 8. A drop in DME price of ĩ0% resulted in a drop in NPV% of ĩ0%, as the profit margin has been decreased due to the decrease in revenue. However, despite this the plant remained profitable displaying the economic strength of the process.

4.3.2 Enterprise discount rate

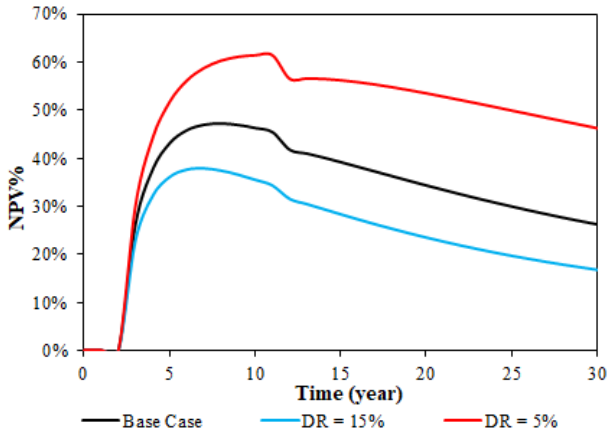


Figure 9: Process economic sensitivity to enterprise discount rate

Figure 9 displays the effect that variations in the enterprise discount rate have on the NPV%. It must be noted that the discount rate assumed does not in fact have any effect on the economic performance of the plant but only on how this performance is measured. Nevertheless, Figure 9 shows to potential investors that even if they discount their investment at a high rate they will still be able to see a return on their investment, thereby establishing the strength of the economic performance of the plant. Therefore, from the sensitivity analysis performed on the economic performance on the designed process, it can be concluded that due to the high profitability of the plant, the process can withstand aggressive changes to the economic model, despite its high sensitivity to these changes.

4.4 Further environmental analysis

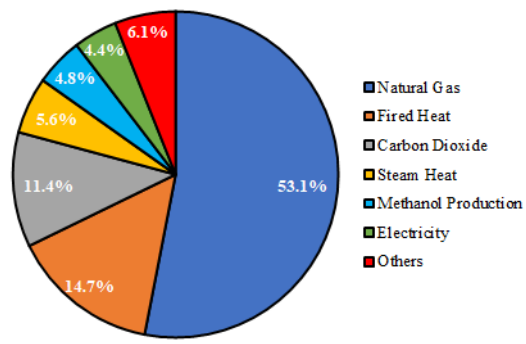


Figure 10: Causes of Environmental Impact

The environmental impact of the selected overall process was analysed to investigate whether it would be possible to reduce the estimated monetary damage of \$2.79/kg_{DME}. The breakdown of the process inputs that are the main contributors to the environmental impact, according to the LCA analysis performed in Simapro, are laid out in Figure 10. From this, it is evident that the majority of the environmental impact of the process is associated with the increasing cost that is incurred due to the increasing scarcity of natural gas as a resource. However, since natural gas is a required feedstock in the production of syngas using a SMR with RWGS production process, there is little that can be done to reduce this impact unless the amount required is reduced. It was analysed whether changing the source of the natural gas from the North Sea to Russia, the Netherlands or Norway could reduce the environmental impact, but these changes in fact did the opposite so this attempt was abandoned. The CO_2 and steam heat markets were assumed to be relatively fixed so unless any waste sources could be sourced, there was little change that could be made to these inputs. Meanwhile, the majority of the impact caused in the production of methanol is due to the high consumption of water in the process for cooling purposes. This is a more environmentally friendly solution than using refrigerants so there is no clear way to reduce this impact without changing the process significantly. However, the fired heat and electricity process inputs can be manipulated to reduce their environmental impact. The fired heater was originally assumed to be a modulating fan boiler with low NO_x emissions, whilst the electricity production input in Simapro, was originally assumed to be from a combined heat and power natural gas plant These were then assumed to be produced from a lean natural gas burner and

wind power, respectively. These changes resulted in a reduction of environmental impact to \$2.49/kg$_{DME}$. Nevertheless the design process selected is still far inferior in this regard to the gasification of biomass process that it was being compared to in Section 4.2, which had an environmental impact of \$1.44/kg$_{DME}$.

5 Conclusion

Based on the initial economic and LCA analysis performed on various DME production processes, it was established that producing DME directly from syngas, whilst more cost effective, had a far greater environmental impact than the indirect process (i.e. via methanol). Therefore, on account of these unacceptable, potential damages to the environment, the indirect production pathways were selected for further economic analysis. During this, it was found producing syngas via the gasification of a biomass slurry, was too expensive a process to allow for the DME production to be economically feasible. This was on account of the high costs of the biomass slurry feedstock. By contrast, producing the initial syngas from natural gas using a SMR with RWGS process allowed the production to be highly profitable. Moreover, due to this high profitability, the economic model was able to remain profitable when subjected to aggressive changes such as a 10% drop in DME price. Hence, this process was concluded to be the overall best performing production route for DME that had been tested. However, the environmental impact of this process was difficult to reduce significantly, remaining significantly poorer than a process using a biomass feedstock to produce the syngas, despite efforts to improve its environmental performance.

6 Outlook

The work in this report could potentially be improved upon in several ways: the process simulations could have used a more rigorous model (e.g. a Gibbs reactor in Aspen HYSYS); simulations for each of the processes could have been performed to allow for fairer comparisons, whilst a midpoint LCA could have been carried out to display precisely where the environmental damage was being caused. Nevertheless, the research found a clear result for an economically feasible DME production process from syngas with an acceptable environmental impact. This result can be built upon in future by comparing it with DME production processes using other feedstocks (e.g. the direct hydrogenation of CO_2) or by analysing the causes of high cost of the biomass slurry to see if they can be reduced in anyway to an economically feasible level.

7 Additional kinetic data

ATR kinetic parameters

Reaction	$k_{0,j}$ (kmol/kg cat h)	E_j (kJ/kmol)
1	$5.852 \times 10^{17} bar^{-1.5}$	204000
2	$4.255 \times 10^{15} bar^{0.5}$	240100
3	$1.955 \times 10^{6} bar^{-1}$	67130
4	$1.020 \times 10^{15} bar^{0.5}$	243900

Table 4: Kinetic parameter

ATR equilibrium constants

Reaction	Equilibrium constant K_{ej}
2	$5.75 \times 10^{12} exp(-\frac{11476}{T}) bar^2$
3	$1.26 \times 10^{-2} exp(\frac{4639}{T})$
4	$7.24 \times 10^{10} exp(-\frac{-21646}{T}) bar^2$

Table 5: equilibrium constant

ATR adsorption constants

Species	$K_{oi}(bar)$	H_i (kJ/kmol)
CH_4 (combustion)	4.02×10^5	103500
O_2 (combustion)	$5.08 \times 10^4 bar^{0.5}$	66200
CH_4	6.65×10^{-4}	-38280
CO	8.23×10^{-5}	-70650
H_2	6.12×10^{-9}	-82900
H_2O	$1.77 \times 10^5 bar$	88680

Table 6: Adsorption constants of ATR

Kinetic parameters of DME production

Parameters	$A(i)exp(\frac{B(i)}{RT})$	
	A(i)	B(i)
k_1	1.65	36696
K_2	3.61×10^3	0
K_3	0.37	17197
K_4	7.14×10^{-11}	124119
k_5	1.09×10^10	-94765
K_{CH_3OH}	7.9×10^{-4}	70500
k_6	3.7×10^{10}	-105000
K_{H_2O}	0.84×10^{-1}	41100

Table 7: DME kinetics parameters

8 Acknowledgements

First, we would like to thank Prof. Chachuat for his continuous support throughout the entirety of our research period. Also, we would like to express our gratitude to Dr. Andrea Bernard, who assisted us in sourcing relevant research articles and designing our Aspen HYSYS simulations. Finally, we would like to thank Mr. Daniel Rodriguez Vallejo for helping us to understand how to best carry out a LCA analysis in Simapro.

9 References

Agrafiotis, C., Rob, M. & Sattler, C. (2015), 'A review on solar thermal syngas production via redox pair-based water/carbon dioxide splitting thermochemical cycles', *Renewable and Sustainable Energy Reviews* **Vol.42**, pp: 254–285.

Air Liquide (2019), 'Autothermal reforming (ATR) - syngas generation', https://www.engineering-airliquide.com/autothermal-reforming-atr-syngas-generation. Accessed: 15/11/2019.

Aldibaba (n.d.), 'Industry grade zno zinc oxide catalyst powderl', https://www.alibaba.com/product-detail/Industry-grade-zno-zinc-oxide-catalyst_724494273.html?spm=a2700.7724857.normalList.16.bcf072b7dmB6Ub&s=p. Accessed: 15/11/2019.

Alternative Fuels Data Center (n.d.), 'Dimethyl ether', https://afdc.energy.gov/fuels/emerging_dme.html. Accessed: 15/11/2019.

American Water Work A (n.d.), 'Is the water free', https://www.pvwc.com/story_of_water/html/costs.htm. Accessed: 15/11/2019.

Baltrusaitis, J. & Luyben, W. L. (2015), 'Methane Conversion to Syngas for Gas-to-Liquids (GTL): Is Sustainable CO2 Reuse via Dry Methane Reforming (DMR) Cost Competitive with SMR and ATR Processes?', *ACS Sustainable Chemistry and Engineering* **Vol.3**, pp: 2100–2111.

Benguerba, Y., Dehimi, L., Virginie, M., Dumas, C. & Ernst, B. (2015), 'Modelling of methane dry reforming over ni/al2o3 catalyst in a fixed-bed catalytic reactor', *Reaction Kinetics, Mechanisms and Catalysis* **Vol.114**, pp: 109–119.

Chibane, L. & Djellouli, B. (2011), 'Methane steam reforming reaction behaviour in a packed bed membrane reactor', *International Journal of Chemical Engineering and Applications* **Vol.2**.

Guillen-Gosalbez, G., Caballero, J. A. & Jimenez, L. (2008), 'Application of life cycle assessment to the structural optimization ofprocess flowsheets', *Industrial and Engineering Chemistry Research* **Vol. 47**, pp:777–789.

Hoang, D. & Chan, S. (2014), 'Modeling of a catalytic autothermal methane reformer for fuel cell applications', *Applied Catalysis A: General* **Vol.268**, pp: 207–216.

Huijbregts, M., Steinmann, Z., Elshout, P., Stam, G., Verones, F., Vieira, M., Hollander, A., Zijp, M. & van Zelm, R. (2016), 'Recipe 2016 v1.1: A harmonized life cycle impact assessment method at midpoint and endpoint level'.

International DME Association (n.d.), 'About dme', https://www.aboutdme.org/index.asp?sid=48. Accessed: 15/11/2019.

Luyben, W. L. (2016), 'Control of parallel dry methane and steam methane reforming processes for fischer–tropschsynbgas', *Journal of Process Control* **Vol.39**, pp: 77–87.

Millipore Sigma (n.d.), 'Aluminum oxide', https://www.sigmaaldrich.com/catalog/product/sigald/199974?lang=en&region=US&cm_sp=Insite-_-noResults_ni/al2o3-_-noResults9-4. Accessed: 15/11/2019.

Misc (n.d.*a*), 'Copper pricel', https://www.macrotrends.net/1476/copper-prices-historical-chart-data. Accessed: 15/11/2019.

Misc (n.d.*b*), 'Iron pricel', http://www.metalary.com/iron-price/. Accessed: 15/11/2019.

Misc (n.d.*c*), 'Nickel', https://hobart.k12.in.us/ksms/PeriodicTable/nickel.htm. Accessed: 15/11/2019.

Ng, K., Chadwick, D. & Toseland, B. (1999), 'Kinetics and modelling of dimethyl ether synthesis from synthesis gas', *Chemical Engineering Science* **Vol.54**, pp: 3587–3592.

Perez-Fortes, M. & Tzimas, E. (2016), 'Techno-economic and environmental evaluation of CO_2 utilisation for fuel production', *JRC science for policy report* .

Pubchem (n.d.), 'Dimethyl ether', https://pubchem.ncbi.nlm.nih.gov/compound/Dimethyl-ethersection=Uses. Accessed: 15/11/2019.

Shen, Y. (2019), 'Techno-economic assessment of direct DME synthesis process'.

Simelys Hernández, M. Amin Farkhondehfal, F. S. M. M. G. S. N. R. (2017), 'Syngas production from electrochemicalreduction of co2: current status and prospective implementation', *The Royal Society of Chemistry 2017* **Vol.19**, pp: 2326–3236.

Suwannapichat, Y., Numpilai, T., Chanlek, N., Faungnawakij, K., Metta Chareonpanich, J. L. & Witoon, T. (2018), 'Direct synthesis of dimethyl ether from co2 hydrogenation over novel hybrid catalysts containing a cue zno zro2 catalyst admixed with wox/al2o3catalysts: Effects of pore size of al2o3support and w loading content', *Energy Conversion and Management* **Vol.159**, pp: 20–29.

Towler, G. & Sinnott, R. (2012), *Chemical Engineering Design, Principles, Practice and Economics of Plant and Process Design, Second Edition*, Elsevier.

Trippe, F., Fröhling, M., Schultmann, F., Stahl, R. & Henrich, E. (2011), 'Techno-economic assessment of gasification as a process step within biomass-to-liquid (BtL) fuel and chemicals production', *Fuel processing Technology* **Vol. 92**, pp: 2169–2184.

Weidema, B. P. (2014), 'Comparing three life cycle impact assessment methods from an endpoint perspective', *Journal of Industrial Ecology* **Vol. 19**, pp: 20–26.

Process Domain and Composition Range Reduction Method for Computer-aided Molecular and Process Design

Tao Liu

Department of Chemical Engineering, Imperial College London, U.K.

Abstract

Integrated computer-aided molecular and process design (CAMPD) guarantees a plant-wide optimum solution of molecular decisions and operating conditions. However, challenges remain in solving mixed-integer nonlinear programming (MINLP) problems due to the high nonlinearities of group contribution (GC) based methods and the presence of implicit design constraints. Three feasibility tests are developed to extend the applicability of process domain reduction methods to the design of aqueous solvent mixtures and to identify a reduced mixture composition range. The proposed tests are applied to the design of a CO_2 chemical absorption process with twelve candidate solvents studied for each test. Statistical associating fluid theory for variable range Mie Potential (SAFT-γ Mie) and GC methods are used for the predictions of mixture properties. Considerable eliminations of infeasible molecular or process design space are achieved in each feasibility test.

Keywords: *computer-aided molecular and process design, feasibility tests, composition range reduction*

1. Introduction & Background

The performance of a chemical process is dependent on many factors, including process topology, unit size, operating conditions and materials utilised in the reaction process. (Eden et al., 2004) Traditionally, computer-aided molecular design (CAMD) has been the key method for identifying desirable processing materials. In CAMD method, the problems of process design and molecular design are usually addressed sequentially (Eden et al., 2004). In other words, the molecular design method is applied with prior quantitative knowledge of the process conditions. However, the performance of processing material is largely dependently on the process conditions. Thus, a sequential approach to decision-making, in which materials are chosen for specific tasks and fixed before the design of the process, can lead to poor performance as it limits the design space and neglects the intrinsic links between molecules and process (Adjiman et al., 2014). To address the interdependency between process conditions and materials, recent researches have given particular attention to integrated computer-aided molecular and process design (CAMPD) in which the problems of process design and molecular design are solved simultaneously. Given that the evaluations of processing materials are conducted across the process design space in CAMPD, the design is guaranteed to arrive at a plant-wide optimum solution.

In general, a CAMPD problem is formulated as mixed integer nonlinear optimisation problem (MINLP). Continuous variables are used to represent continuous process conditions such as temperature, pressure and mixture composition (Papadopoulos et al., 2017). On the other hand, discrete variables are used to represent the chemical structure of materials constructed using group contribution (GC) based methods (Odele and Macchietto, 1993).

Over the past few decades, various MINLP solutions algorithms have been developed, namely the outer-approximation (OA) algorithm (Duran and Grossmann, 1986) and the generalized Benders decomposition (Geoffrion, 1972) in which the original MINLP is decomposed into a continuous non-linear problem (NLP) and a mixed integer linear problem (MILP) and they are solved iteratively. Despite the advances that have been made in tackling MINLP, most of group contribution methods and statistical associating fluid theory equations of state (SAFT EoSs) are non-convex and non-linear (Papaioannou et al., 2014), making it difficult for local optimisation solvers to converge to a global optimum (Gopinath et al., 2016). In addition, challenge remains in providing a good initial guess for solvent structures, mixture compositions and corresponding process

conditions to enable the solution of CAMPD problem. Furthermore, there exists molecules that satisfy all the molecular design constraints but violate implicit process constraints. Taking the melting point of a solvent in carbon capture process as an example, one may conclude that the process is infeasible if the operating temperature is below the melting point of the solvent since solvent will be in solid state. Other more complicated process constraints such as phase coexistence and stream purity targets also present, and violation of these constraints will render the process to be infeasible (Gopinath et al., 2016).

To address these issues in solving CAMPD problem, many solution algorithms have recently been developed. Buxtons et al. (1999) was one of the first who embedded molecule property tests to form a subset of CAMD problem constraints and mass-transfer feasibility tests of the process. In Gionavaglou et al.'s (2003) work, similar physical property tests, including boiling point test, phase split test and density test was adopted in a mixed-integer dynamic optimisation (MIDO) framework for solvent design. A more recent hierarchical optimisation (HiOpt) methodology to the solution of CAMPD was suggested by Burger et al. (2015) The method is applied to the design of a physical absorption process. In this case, a reduced process model is first used for an initial multi-objective optimisation (MOO) to obtain a good initial guess for the original design problem. Although the solution of CAMPD becomes numerically less challenging as an initial point is now available, the HiOpt algorithm needs to be developed to accommodate larger search spaces for solvent design and to allow the design of solvent mixtures. Most recently, a modified outer-approximation (MOA) algorithm has been proposed by Gopinath et al. (2016). The MOA algorithm followed a similar feasibility test approach first deployed by Buxton et al. (1999). However, the MOA algorithm differed from previous methods in the sense that the process feasibility is assessed for combinations of molecular and process variables. Such feasibility test can, therefore, not only determine the feasibility of a solvent in a process, but also provide a reduced process domain for a solvent once the solvent is deemed feasible. Although the algorithm is demonstrated to be successful for a physical absorption process to separate carbon dioxide from methane, the proposed MOA approach remains only applicable to the design of pure solvent.

Building as an extension of MOA approach proposed by Gopinath et al. (2016), this work aims to develop several feasibility tests for CAMPD problem of a post-combustion CO_2 chemical absorption process. The tests differ from previous work primarily in a way that a mixture of solvent and water is now considered instead of pure solvent. This will greatly increase the complexity of the design problem since the solvent mixture composition must be considered as an additional continuous variable. Thus, in addition to the identification of solvent feasibility and reduced process domain, the proposed test will also provide a reduced solvent mixture composition domain in which the process remains feasible. Lastly, the work aims to demonstrate the combined use of SAFT-γ Mie EoS and other group contribution methods in formulating the feasibility tests.

2. Methodology

2.1 Process Description

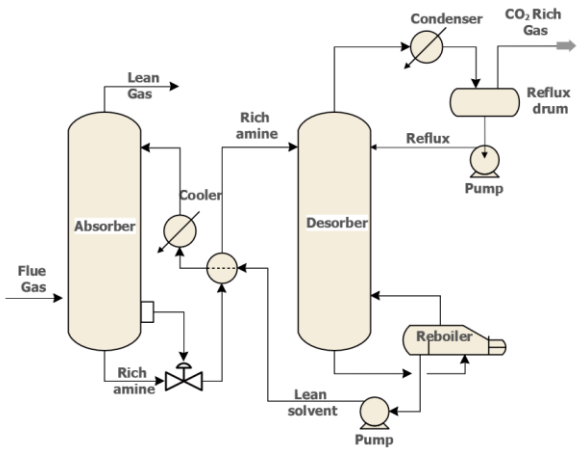


Figure 1. A general flow sheet for amine-based CO_2 chemical absorption process

To better illustrate the proposed feasibility tests, a simplified post-combustion CO_2 chemical absorption process presented in Figure 1 is studied. The two major process units in the flowsheet are the counter-current absorber and desorber. A post-combustion flue gas that usually contains carbon dioxide, nitrogen and other impurities is fed to process from the bottom of the absorber while lean

aqueous solvent mixture enter the absorber from the top to extract CO_2 from the flue gas stream. A lean gas outlet stream containing mainly nitrogen exits from the top of absorber column while a CO_2-rich solvent stream is obtained at the bottom and fed to the desorber column. The simplified desorber column contains two additional subunits including a condenser and a reboiler. The CO_2-rich solvent stream entering the desorber is separate into two streams with a CO_2-rich gas stream produced at the top and the solvent mixture recovered at the bottom. As a recycle stream, the recovered lean solvent is then fed back to the absorber column for further carbon dioxide capture.

2.2 Modelling Assumption

Due to the complexity of the absorption process, the following assumption is made so that trackable feasibility test can be developed:

The flue gas stream is assumed to contain only carbon dioxide and nitrogen. For a post-combustion capture process, the composition of CO_2 in the flue gas steam typically accounts 5% - 15% (Leung, Caramanna and Maroto-Valer, 2014), depending on the fuel being used, with most of the rest being N2. Although many other impurities such as nitrogen oxides, sulphur oxides and oxygen also exist in the flue gas, the amount of the impurities is negligible when compare with the abundance of carbon and nitrogen present in the stream. Furthermore, the SAFT-γ Mie EoS employed in this work is not capable of accounting for the degradation of amine in the presence of O_2, CO_2, NOx, SO_2. Thus, removal of impurities from the flue gas stream will have a minimal effect to test results and design solutions.

2.3 Thermodynamic Property Predictions

With recent advances in GC-based methods, in particular the development of SAFT equations of state, a wide range of thermodynamic properties for solvents and their mixtures can be predicted more reliably (Lymperiadis et al., 2007). In Gopinath et al.'s (2016) work, most thermodynamic properties were predicted using SAFT-γ Mie EoS. Similar as other GC approaches, SAFT-γ Mie EoS predicted molecular properties by dividing a molecule into molecular descriptors and summing the contribution of each descriptor to obtain the thermodynamic properties of a molecule. However, a SAFT-γ Mie EoS utilises a fused heteronuclear model in which the interactions between monomeric segments are represented through Mie potentials, a generalized Lennard-Jones potential, of variable attractive and repulsive range. Such treatment of segment interactions allows better description of the fluid-phase equilibria of real compounds (Papaioannou et al., 2014). In this work, SAFT-γ Mie EoS is chosen to predict solvent boiling points and equilibrium compositions of mixtures in the desorber.

Melting points of pure components are predicted using GC approach proposed by Hukkerikar et al. (2013)

$$\exp\left(\frac{T_m}{T_{m0}}\right) = \sum_i n_i a_i \qquad (1)$$

where T_m(K) is the normal boiling point of a molecule and T_{m0}(K) is the additional adjustment parameter for the prediction model. n_i and a_i represents the occurrence and contribution of molecular group i respectively.

Viscosities of solvent mixtures are predicted through the integration of GC method proposed by Hsu and Li (1997) and linear mixing rule. Viscosities of pure solvents are calculated using the following equation:

$$\ln \eta = \sum_i n_i \{a_i + b_i T + c_i T^{-2} + d_i \ln P_c\} \qquad (2)$$

where η(mPas) is the viscosity and T(K) is the temperature of a solvent. a_i, b_i, c_i and d_i are the predictive constants of group i. P_c (bar) represents the critical pressure of a solvent and it is obtained using the critical property prediction method proposed by Gani et al. (2001) The function of P_c(bar) is given as

$$(P_c - P_{c1})^{-0.5} - P_{c2} = \sum_i n_i p_i \qquad (3)$$

where P_{c1} (bar) and P_{c2} (bar) are additional adjustment parameters for critical pressure prediction and p_i is the contribution of group i.

2.4 Feasibility tests

For decomposition based MINLP solution methods, a new candidate solvent is generated at each major

iteration of MILP, the master problem. An NLP, the primal problem, is then formulated by fixing integer variables to the solution of the master problem. However, the generated candidate solvent from the master problem can often lead to numerical failures in the solution of the primal problem. This is because there exist infeasible regions for the continuous variables in which implicit process design constraints are violated. In this work, three tests are developed to examine solutions of the master problem against several implicit process constraints. By implementing the proposed tests, infeasible solvents can be removed from the design search space before they are used in the solution of the primal problem. In addition, reduced process domains and composition ranges can be identified through the tests. The infeasible domains caused by implicit process constrains can thus be avoided. Lastly, the reduced process domains can also provide a good initial guess for the primal problem since any points within these domains will lead to the convergence of the primal problem.

Table 1. Candidate solvents used in tests and their molecular structures

Category	Solvent Name	Molecular Structure
Primary	MEA	$CH_2NH_2CH_2OH$
Primary	AMP	$CNH_2(CH_3)_2CH_2OH$
Tertiary	MDEA	$CH_3N(CH_2)_4OH$
Tertiary	TEA	$CH_2N(CH_2)_5(OH)_3$
Primary	Ethylamine	$CH_3CH_2NH_2$
Tertiary	Trimethylamine	$CH_3N(CH_3)_2$
Primary	Candidate 1	$CNH_2(CH_2)_2(OH)_3$
Primary	Candidate 2	$CNH_2(CH_3)_4CCH_2OH$
Primary	Candidate 3	$CHNH_2(CH)_2CH_2(OH)_3$
Tertiary	Candidate 4	$CH_3N(CH_2)_6OH$
Tertiary	Candidate 5	$CH_3N(CH_2)_8OH$
Tertiary	Candidate 6	$CH_3N(CH_2)_4OHCH_3$

For each test proposed in this work, six primary amines and six tertiary amines are studied to provide a better understanding of effectiveness of the tests for different type of amines. Names and molecular structures of the candidate amines are provided in Table 1. In test 1, the melting points of pure amine is examined to determine whether the solvent is in liquid state under ambient conditions. Test 1 differs from the rest in the sense that it solely examines the property of pure components while other tests are carried out for mixtures. As a result, solvent will be eliminated from the search space if test 1 is failed. In test 2, the boiling points and viscosities of water-solvent mixtures are tested for all possible compositions. The test results provide reduced composition ranges in which the boiling point and viscosity of a solvent satisfy the implicit process constraints embedded in this stage. Test 3 aims to identify an upper bound of the desorber pressure beyond which the lean solvent cannot meet the desired purity target.

2.4.1 Test 1: Handling feasibility test for solvent melting point

In test 1, a user-specified solvent handling temperature T_s (K) is first defined. This corresponds to the temperature of solvent mixture that enters the absorber as a fresh feed. For an amine-based CO_2 absorption process, the operating pressure in the absorber column is normally around the atmospheric pressure. Thus, the pressure is fixed at 1.013 bar in our case. For solvent handling to be feasible, a solvent mixture must be in a liquid state at the given operating conditions. Therefore, the melting point of a solvent mixture should be higher than T_s:

$$T_{melting}\left(P = 1\ atm, n^j\right) > T_s \qquad (4)$$

where $T_{melting}$(K) represents the melting point of the mixture and n^k stands for candidate solvent j.

In general, the $T_{melting}$ of a mixture varies with changes in component compositions and a nonlinear relationship between the two properties can be observed, making a linear mixing rule inapplicable. Since all available GC methods cannot predict $T_{melting}$ for a wide range of amines, an alternative screening approach proposed by Papadopoulos et al. (2017) has been adopted. In Papadopoulos' work on examining the melting point of solvent mixture, a simple assumption has been made that a solvent mixture is feasible if the highest melting point of the one of the two mixture components is higher than the desired handling temperature. Such assumption can be regarded as an overestimation of melting point when amine and water are considered as the two components of the mixture. In this case, the melting point of water is well below T_s while amines can vary across a large range around T_s. As a result, the melting point of

an amine-water mixture is likely to be lower than the maximum of mixture components. To account for the overestimation of melting point, a temperature allowance α(K) is proposed to loosen the upper bound of T_s. The test formulation is given as:

$$T_{melting}(P = 1\ atm, n^j) > T_s + \alpha \qquad (5)$$

2.4.2 Test 2: Handling feasibility test for solvent boiling point and viscosity

Prior to test 2, a user-specified maximum temperature T_s^b(K) is defined for the solvent inlet stream. In this work, T_s^b is set as 313.15K by taking the process parameters at the UKCCSRCPACT pilot plant as references (Bui et al., 2018). Similar as in test 1, solvent handling is feasible only if the solvent mixture is in liquid state. This is only satisfied if the boiling point of mixture is higher than T_s^b. The test for boiling point is formulated through the following equations:

$$T^b(n^j, x_{H_2O}^k) > T_s^b \ \forall\ x_{H_2O}^k \qquad (6)$$

$$0 < x_{H_2O}^k < 1 \qquad (7)$$

where T^b(K) is the normal boiling point of a solvent mixture and $x_{H_2O}^k$ is a continuous variable that corresponds to the water mass fraction of the mixture k. The feasibility of a solvent mixture is examined at all compositions and one or more reduced compositions range that satisfy the boiling point constraint can then be obtained. If a solvent mixture is infeasible at all compositions, it is subsequently removed from the design search space.

Additional test is proposed for solvent viscosity v(mPas) since the maximum viscosity v_{max}(mPas) that can be handled in an absorption process is limited by the pump power. In this work, a common v_{max} of 100 $mPas$ is used as the test criterion. For GC property prediction methods, similar limitation also exists in the predicting viscosities of mixtures as it does in predicting melting points. Thus, a linear mixing rule is adopted to calculate the viscosity of a mixture such that:

$$v_{mix} = x^k v_{H2O} + (1 - x^k) v_k \qquad (8)$$

where v_{H2O}(mPas) and v_k(mPas) stand for the viscosity of pure water and pure solvent k respectively. Although such assumption may not seem fully justified in real case, the primary aim of this work is to demonstrate a framework under which GC based methods can be applied to predict viscosities of a solvent mixture. The test of viscosity is formulated in the same manner as that of boiling point such that solvent feasibility is tested across all compositions of mixture. As a result of the test, a solvent is either removed from the design search design if it is infeasible at all compositions or provided with reduced feasible composition domains. The following equations are given as the formulation of the viscosity feasibility test:

$$v_{mix}(n^j, x^k) > v_{max} \ \forall\ x^k \qquad (9)$$

$$0 < x^k < 1 \qquad (10)$$

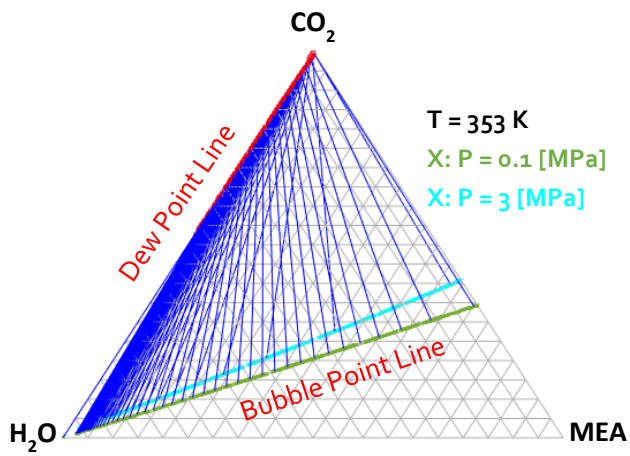


Figure 2. A ternary phase diagram for CO_2-H_2O-MEA at T=353K and P = 0.1 and 3 MPs

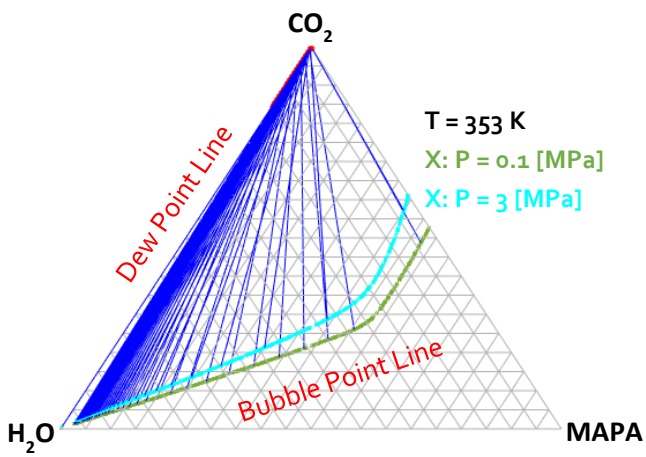


Figure 3. A ternary phase diagram for CO_2-H_2O-MAPA at T=353K and P = 0.1 and 3 MPs

2.4.3 Test 3: Feasibility test for the CO2 loading in desorber lean solvent stream

The design of carbon dioxide desorber usually involves setting the CO_2 loading of the lean solvent stream so that the recycled solvent contains a reasonable amount of carbon dioxide. Additionally, the CO_2 loading, also known as lean loading, has a significant effect on the heat duty of reboiler. A regeneration study carried out by Oexmann and Kather (2010) reported that a lean loading of 10% to 20% can be achieved at a relatively low heat duty when 30%wt monoethanolamine (MEA) is used as solvent. To account for the implicit lean loading constraint, test 3 is used to determine an upper bound P^u (bar) of the operating pressure at the bottom stage of the desorber beyond which the desired CO_2 loading is thermodynamically unattainable. In other words, under given process conditions, there must exist an equilibrium point of the liquid phase at which the CO_2 composition is less than the required leaning loading.

In test 3, it is assumed that the mole fraction of CO_2 in liquid phase increases monotonically as the operating pressure, P (bar) increases. Such assumption allows a maximum pressure P_{max}(bar) to be identified since any pressure beyond this point will result in an even greater minimum lean loading. Figure 2 and 3 present the ternary diagrams of the phase envelope for the mixture of amine, carbon dioxide and water. By increasing the pressure from 0.1 to 3 MPa while keeping temperature the same, it is observed that the two-phase envelope becomes smaller. As a result, the minimum CO_2 fraction in the liquid phase also decreases. Since the relationship is true for both MEA and MAPA, the assumption of monotonicity is believed to be justified.

Additionally, no limit is posed for the number of stages that the desorber can have. Although it may be impractical in real process design, such assumption allows the purification feasibility test in this work to identify the highest attainable purity target for the lean solvent stream.

The bottom stage of desorber is assumed to be in equilibrium since the feasibility test for the desorber makes use of equilibrium flash algorithm in gPROMS modelling environment to examine the theoretical compositions in the bottom stage.

Thus, test 3 is formulated as an optimisation problem to identify a maximum operating pressure of the desorber. The optimisation formulation is given as follows:

$$P_{max} = \max_{P,T,x^k,x_{H_2O},x_{CO_2},y^k,y_{H_2O},y_{CO_2}} P \quad (11)$$

s.t.

$$\mu_i^V(P,\ T, y^k, y_{H_2O}, y_{CO_2}, n^j) = \mu_i^L(P,\ T, x^k, x_{H_2O}, x_{CO_2} n^j) \quad (12)$$

$$0 < N^V \quad (13)$$

$$0 < N^L \quad (14)$$

$$x^k + x_{H_2O} + x_{CO_2} = 1\ \forall\, x^k \quad (15)$$

$$y^k + y_{H_2O} + y_{CO_2} = 1\ \forall\, y^k \quad (16)$$

$$P^L < P < P^U \quad (17)$$

$$T_s + \alpha < T < T^U \quad (18)$$

$$0 < x^k < 1 \quad (7)$$

$$x_{H_2O}^L < x_{H_2O} < 1 \quad (19)$$

$$\frac{x_{CO_2}}{x^k} < x^{loading} \quad (20)$$

The first constraint ensures that a chemical equilibrium is reached on the bottom stage of the reboiler. μ_i^V (J/mol) and μ_i^L (J/mol) represent the chemical potentials in the vapour and liquid phase while y and x represent the compositions of each phase. In constraint 2 and 3, N^V and N^L stand for the number of molecules in the vapour and liquid phase respectively. Constraining both N^V and N^L to be greater than 0 ensures the co-existence of vapour and liquid phase. The operating pressure of the desorber is bounded between 1 and 100 bar. As seen in Figure 2 and 3, the lowest attainable leaning loading increases as the mixture composition moves along the bubble point line to the left. Although the presence such points help expand the feasible domain of the desorber operating pressure, the compositions of these points do not reflect that of a typical lean solvent stream. As the composition move to left of the ternary diagram, the water fraction of liquid phase decreases. In reality, the water fraction of lean solvent stream is usually much higher the fraction of other components. Therefore, a lower bound of water fraction x_{H2O}^L is set in test 3. Lastly, the constraint on ratio between CO_2 and solvent ensures that the equilibrium composition of the liquid phase has a CO_2 fraction

lower than the required CO_2 loading. In our work, $x^{loading}$ is set to be 0.1.

3. Result and Discussion

3.1 Test 1: Handling feasibility test for solvent melting point

To study the effect of temperature allowance on test results, twelve candidate solvents are tested with α taking values of 0 and $50K$. As shown in Figure 2, only one-third of the twelve solvents, including MEA, ethyl amine, trimethylamine and candidate 6, has passed the test when α is set to be 0. However, it is important to note that several commonly used CO_2 capture solvents such as MDEA, TEA, and AMP have failed the melting point feasibility test. This can be partially attributed to the overestimation of melting points using GC method. A comparison between GC predictions and experimental values revealed that the errors in the model prediction become significantly large as the complexity of molecular structure increases. Although the applied GC method is very accurate in predicting the melting points of primary amines such as MEA and trimethylamine, predictions for all tertiary amines including MDEA, TEA and trimethylamine have deviations of more than 30K from their corresponding literature values. As a result, MDEA and TEA are screened out by test 1 although both solvents have been widely proved feasible.

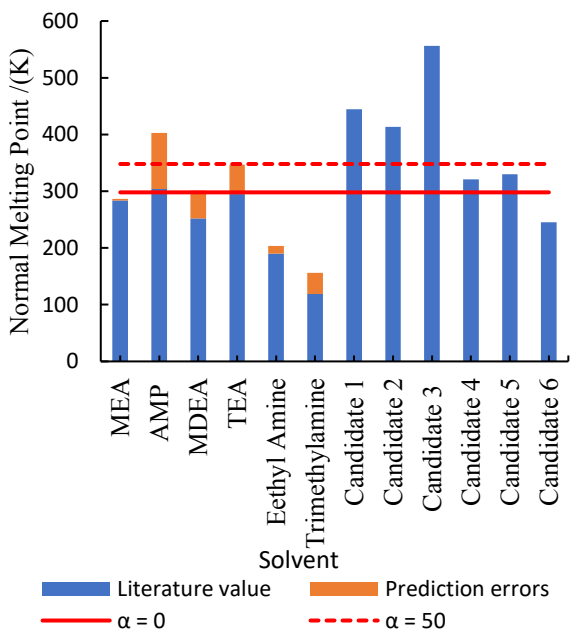


Figure 4. Solvent melting point test results

Additionally, overestimation is also led by the fact that test 1 examines highest melting point of components in a solvent mixture. For instance, although experimental result shows that melting point of AMP is slightly above T_s, a mixture of water and AMP may be feasible as melting point can vary largely with change in the composition of mixtures.

On the other hand, four additional solvents, including MDEA, TEA, candidate 4 and 5, passed the test when α is raised to $50K$. In this case, AMP is the only known feasible solvent that failed the test due the significant overestimation of its melting point.

Overall, out of six randomly generated amine solvents, test 1 is able identify three as infeasible solvents. Such results suggest that a melting point test can be effectively used to eliminate infeasible molecules from the design space. Furthermore, it is notable that test results can be greatly affected the accuracy of model predictions and assumptions in test formulation. Thus, a suitable temperature allowance should be incorporated with the test to mitigate the overestimation of solvent melting point.

3.2 Test 2: Handling feasibility test for solvent boiling point and viscosity

Despite the elimination of some solvents in test 1, all of the twelve candidate solvents are tested again in test 2. Table 2 and 3 draws the comparison between model predictions and literature values of both boiling points and viscosity of pure components.

Table 2. Literature values and model predictions of boiling points of pure solvents at T = 298K

Solvent Name	Literature Value (K)	Model Prediction (K)	% Error
MEA	443	442.83	0.04
AMP	438	437.33	0.29
MDEA	520	501.44	3.61
TEA	608	544.28	10.56
Ethylamine	290	290.08	0.03
Trimethyl-amine	278	270.15	2.88

Table 3. Literature values and model predictions of solvent viscosity of pure solvents at T = 298K

Solvent Name	Literature Value (K)	Model Prediction (K)	% Error
MEA	17.90	5.56	62.46
AMP	n/a	10.86	n/a
MDEA	77.19	69.95	1.68
TEA	662.5	1167.05	11.13
Ethylamine	0.28	0.27	0.03
Trimethyl-amine	0.18	0.17	2.88

The results show that SAFT-γ Mie EoS has a high level of accuracy in terms of boiling point prediction. Except only for TEA, the predictions for all other known solvents exhibit deviations of less than 5% from literature values. In contrast, significant errors are present in the predictions of viscosities obtained through other GC methods. The percentage error of model prediction is found at its greatest, 62.46%, in the case of MEA.

Table 4. Literature values and model predictions of solvent critical pressure of pure solvents

Solvent Name	Literature Value (K)	Model Prediction (K)	% Error
MEA	80.30	66.48	17.21
AMP	n/a	42.64	n/a
MDEA	41.6	39.18	5.83
TEA	34.82	34.28	1.55
Ethylamine	56.93	57.71	1.37
Trimethyl-amine	40.72	41.66	2.31

However, such inaccuracy should be attributed not only to the prediction method of viscosity, but also to that of critical pressure. Results in Table 4 has shown that the percentage error of P_c is 17.21% for MEA. Since viscosity is a function of molecular structure, temperature and critical pressure, an error in P_c can seriously affect the accuracy of model prediction.

Figure 5 presents the results of boiling point test for solvent mixtures. Infeasible composition ranges are found for ethylamine and trimethylamine whereas the rest of candidate solvents remains feasible throughout the composition spectrum. Prior to test 2, no monotonicity between boiling point and mixture composition was assumed, that is, more than one infeasible region may be identified. However, it is found that the boiling points of all twelve mixtures always change monotonically as the fraction of solvents increases. Ethylamine is found to be infeasible once the solvent fraction goes beyond 70.6% whereas a relatively lower solvent fraction of 26.7% was found for trimethylamine. This indicates that test 2 may be applied exclusively to low-boiling solvents ($T^b < T_s^b$) such as ethylamine and trimethylamine to identify a maximum solvent fraction in the mixture since other high-boiling solvents can always pass the test.

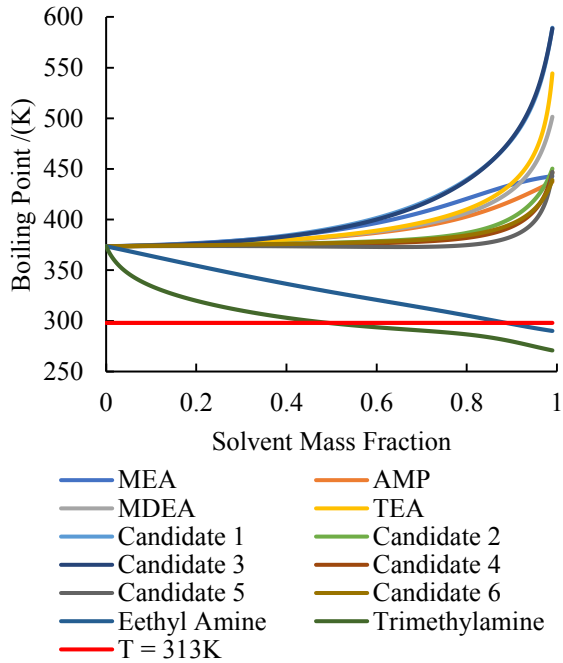


Figure 5. Solvent boiling point test results

Since a linear mixing rule is assumed when calculating the viscosity of a solvent mixture, monotonicity between viscosity and solvent mass fraction is also observed from the test result shown in Figure 6. In the case when v_{max} is equal to 100 $mPas$, maximum solvent fractions were found to be 43.3% for TEA, 8.5% for candidate solvent 1 and 1.5% for candidate solvent 3 whereas the rest are found to be feasible with all mixture compositions. To avoid elimination of feasible composition range as a result of errors in viscosity prediction, test was also performed with v_{max} set to 150 $mPas$ and significant increase of maximum solvent fractions to 54.7% is observed for TEA while changes for candidate solvent 1 and 3 are minimal.

Overall, the result of test 2 has suggested that a feasibility test based on solvent mixture boiling

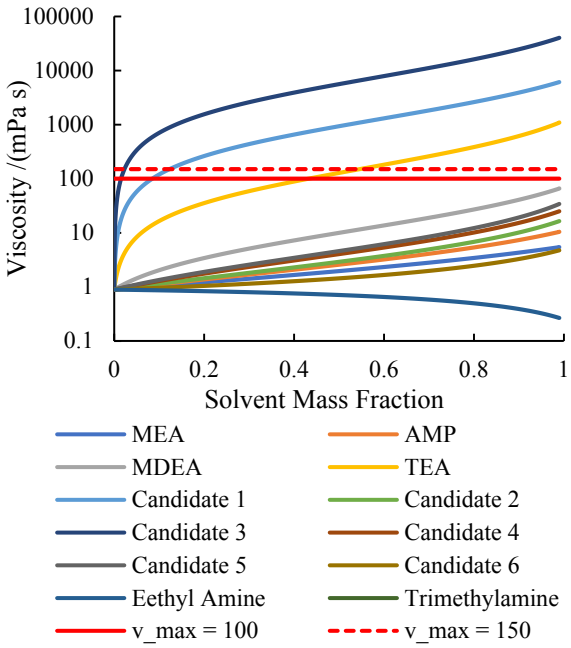


Figure 6. Solvent viscosity test results

points can reliably produce an upper bound of solvent fraction when such method is coupled with SAFT-γ Mie EoS. In comparison, the model predictions of viscosity exhibit a large discrepancy from literature values. To avoid elimination of feasible mixture compositions, setting v_{max} at reasonably higher level is found essential in accounting for errors in model predictions.

3.3 Test 3: Feasibility test for the CO2 loading in desorber lean solvent stream

Test 3 was carried out with the molar fraction of water in the lean solvent stream set to be greater than 89%. A summary of P_{max} and active constraint for each candidate solvent is presented in Table 5. Except for ethylamine and trimethylamine, test 3 is successful in identifying a reduced operating pressure domain for all candidate solvents whereas ethylamine and trimethylamine are found feasible across the tested operating pressure range. Comparisons between the pre-defined and the reduced domains show that the test has potential of eliminating more than 75% of pressure domain which is infeasible.

In addition, distribution patterns are observed such that all of the primary amines except ethylamine have P_{max} ranging around 10 bar whereas P_{max} for all tertiary amines except trimethylamines was found near 20 bar. Such result is consistent with previous findings in literature that primary amines are generally found with better CO_2 affinity than higher-order amines. Given the assumption that lean loading increases as the operating pressure increases, P_{max} is likely to be lower for primary amines since they can react more readily with CO_2.

Table 5. Outcome of test 3 and active constraints of optimisation

Solvent Name	P_{max} (bar)	Active Constraint	% Domain Reduction
MEA	10.22	$x^{loading}$	89.67
AMP	10.20	$x^{loading}$	89.70
MDEA	21.80	$x^{loading}$	77.98
TEA	22.16	$x^{loading}$	78.62
Ethylamine	100.00	P^U	0
Trimethylamine	100.00	P^U	0
Candidate 1	10.26	$x^{loading}$	89.64
Candidate 2	10.17	$x^{loading}$	90.74
Candidate 3	10.28	$x^{loading}$	89.62
Candidate 4	18.46	$x^{loading}$	81.35
Candidate 5	19.94	$x^{loading}$	79.86
Candidate 6	19.04	$x^{loading}$	80.77

4. Conclusions

In test 1, melting points of pure solvents are examined and some conventional CO_2 absorption solvents including MDEA, TEA and AMP are found infeasible due to significant errors in model predictions. Incorporation of suitable temperature allowance within the test is proved to generate a much more reliable test result, leaving AMP as the only known feasible solvents that fails the test.

In test 2, two additional tests are implemented to screen aqueous solvent mixtures based on boiling points and viscosities. The results suggest that such tests can be effectively used in identifying a reduced feasible composition range. However, adjustments to the test criteria for solvent viscosities are found necessary to account for model prediction errors.

In test 3, the feasibility of lean loading target is examined to obtain a reduced operating pressure domain of the absorber. With the incorporation of constraint on the water fraction in the lean solvent stream, the test is successful in identifying a reduced operating pressure domain within which a

practical composition of lean solvent can also be reached.

Overall, the proposed feasibility tests are proved to be successful in identifying reduced process domain and mixture composition range. Future work is suggested to implement the feasibility tests to a wide range of CAMPD problems so that the performance of the tests can be assessed with a large molecular and process design space. In addition, significant errors in GC model predictions are identified. To improve the reliabilities of test results, better prediction methods for mixture properties such melting points and boiling points should be developed.

5. Acknowledgements

The author would like to thank PhD supervisor, Lauren Lee for her guidance along the research project and undergraduate student, Yu Zhang for sharing her expertise in gPROMS modelling.

6. References

Adjiman, C., Galindo, A. and Jackson, G. (2014). Molecules Matter: The Expanding Envelope of Process Design. *Proceedings of the 8th International Conference on Foundations of Computer-Aided Process Design*, pp.55-64.

Buxton, A., Livingston, A. and Pistikopoulos, E. (1999). Optimal design of solvent blends for environmental impact minimization. *AIChE Journal*, 45(4), pp.817-843.

Bui, M., Tait, P., Lucquiaud, M. and Mac Dowell, N. (2018). Dynamic operation and modelling of amine-based CO2 capture at pilot scale. *International Journal of Greenhouse Gas Control*, 79, pp.134-153.

Burger, J., Papaioannou, V., Gopinath, S., Jackson, G., Galindo, A. and Adjiman, C. (2015). A hierarchical method to integrated solvent and process design of physical CO2absorption using the SAFT-γ Mie approach. *AIChE Journal*, 61(10), pp.3249-3269.

Duran, M. and Grossmann, I. (1986). An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming*, 39(3), pp.337-337.

Eden, M., Jørgensen, S., Gani, R. and El-Halwagi, M. (2004). A novel framework for simultaneous separation process and product design. *Chemical Engineering and Processing: Process Intensification*, 43(5), pp.595-608.

Geoffrion, A. (1972). Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, 10(4), pp.237-260.

Giovanoglou, A., Barlatier, J., Adjiman, C., Pistikopoulos, E. and Cordiner, J. (2003). Optimal solvent design for batch separation based on economic performance. *AIChE Journal*, 49(12), pp.3095-3109.

Gopinath, S., Jackson, G., Galindo, A. and Adjiman, C. (2016). Outer approximation algorithm with physical domain reduction for computer-aided molecular and separation process design. *AIChE Journal*, 62(9), pp.3484-3504.

Hukkerikar, A.S., Sin, G., Abildskov, J., Sarup, B. and Gani, R. (2013). Development of pure component property models for chemical product-process design and analysis. *Technical University of Denmark, Department of Chemical and Biochemical Engineering, Kgs. Lyngby. Chemical Physics*, 127(23), p.234903.

Hsu, C. and Li, M. (1997). Viscosities of Aqueous Blended Amines. *Journal of Chemical & Engineering Data*, 42(4), pp.714-720.

Lymperiadis, A., Adjiman, C., Galindo, A. and Jackson, G. (2007). A group contribution method for associating chain molecules based on the statistical associating fluid theory (SAFT-γ). *The Journal of Chemical Physics*, 127(23), p.234903.

Leung, D., Caramanna, G. and Maroto-Valer, M. (2014). An overview of current status of carbon dioxide capture and storage technologies. *Renewable and Sustainable Energy Reviews*, 39, pp.426-443.

Marrero, J. and Gani, R. (2001). Group-contribution based estimation of pure component properties. *Fluid Phase Equilibria*, 183-184, pp.183-208.

Oexmann, J. and Kather, A. (2010). Minimising the regeneration heat duty of post-combustion CO2 capture by wet chemical absorption: The misguided focus on low heat of absorption solvents. *International Journal of Greenhouse Gas Control*, 4(1), pp.36-43.

Odele, O. and Macchietto, S. (1993). Computer aided molecular design: a novel method for optimal solvent selection. *Fluid Phase Equilibria*, 82, pp.47-54.

Papadopoulos, Athanasios I., Seferlis, Panagiotis, Editor, and Wiley InterScience. Process Systems and Materials for CO2 Capture: Modelling, Design, Control and Integration (2017).

Papaioannou, V., Lafitte, T., Avendaño, C., Adjiman, C., Jackson, G., Müller, E. and Galindo, A. (2014). Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments. *The Journal of Chemical Physics*, 140(5), p.054107.

Towards the Development of a Computer-Aided Molecular and Process Design Framework for Mixtures as Organic Rankine Cycles Working Fluids

Miriam Sarkis and Natsuko Robert
Department of Chemical Engineering, Imperial College London, U.K.

Abstract

Organic Rankine Cycles (ORCs) offer a route to tackle today's environmental and energy challenges by valorising heat recovery from low temperature sources. Optimisation frameworks for pure working fluids have been widely researched. However, the use of mixtures as ORC working fluids has the potential to further increase the thermo-economic efficiency of the cycles. This paper presents the initial steps to develop a novel outer-approximation CAMPD algorithm augmented with feasibility tests using the SAFT-γ Mie group contribution method for accurate property prediction. This work primarily addresses the ORC design primal problem, which identifies optimal process conditions and mixture compositions for a binary combinations of fluids. The proposed feasibility tests discard infeasible regions from the search space, aiding convergence to the global optimum. Given heat source specifications typical of geothermal applications, the tests and primal problem yield results that highlight the limitations of heuristics-based performance predictions for mixtures. For some cases, mixing top performing fluids can lead to a lower power output; conversely, introducing a less efficient or infeasible compound can increase the power generation. The absence of a defined trend thus underlines the necessity of a full CAMPD algorithm to research promising mixtures.

1 Introduction

Growing concerns regarding climate change and energy prices have highlighted the need to recover energy from alternative and more sustainable sources while remaining economically competitive to fossil fuels. Organic Rankine Cycles (ORC) offer a promising and mature route to address this challenge, allowing the conversion of heat from renewable resources to electricity. Operating on the same basis as conventional steam cycles, the use of an organic working fluid with a lower boiling point than water allows the capture of heat at lower temperatures, ranging between 100-350°C (Zhang et al., 2016).

The large spectrum of fluid choices allows flexible and more efficient heat recovery, valorising power generation from biomass as well as conversion of geothermal heat and solar concentration (Markides, 2015). ORCs have also demonstrated potential to reduce the energy intensity of industrial processes and power plants, recovering heat at exhaust temperatures at which steam cycles would be uneconomical due to low efficiency (Macchi et al., 2017). The thermal and economic efficiency of ORC systems for a given heat source is, however, highly dependant on the interlink between the intrinsic thermodynamic behaviour of the working fluid and the operating conditions of the process (Linke et al., 2015).

Holistic approaches for the determination of the best ORC design and working fluid are thus necessary. To that end, research has focused on the implementation of Computer-Aided Molecular and Process Design (CAMPD) algorithms, which address the interdependence of design decisions by simultaneously optimising working fluid structure and ORC process conditions.

2 Background

2.1 Literature Review: Pure fluids

Extended research has been carried out to identify the optimal working fluid and select the corresponding optimal operating conditions.

Papadopoulous et al. (2010) pioneered the use of a Computer-Aided Molecular Design (CAMD) approach for the selection of ORC working fluids. A list of promising candidate molecules was compiled, considering economic, safety and environmental criteria. Properties of interest were computed via Group Contribution (GC) methods and cubic equations of state (EoS) and compared to heuristics-based targets. The performance of the synthesised compounds in the ORC process would be evaluated in a separate stage. The work of Schilling et al. (2017) addressed the limitations of GC prediction methods in accurately predicting fluid phase equilibrium, relying on PC SAFT as an EoS. The CAMPD formulation introduced increased the solution's robustness by yielding the working fluid structure and optimal process conditions in one stage.

White et al. (2018) implemented the state-of-the-art version of SAFT, the SAFT-γ Mie GC framework, with an outer approximation (OA) algorithm to identify the global optimum to the ORC problem. The selection

of compounds was however restricted to small families of hydrocarbons, which limited the design space. The concept of feasibility tests introduced by Bowskill et al. (2016) allowed the exploration of a vaster feasible region, with faster convergence to the solution by discarding infeasible fluids and guiding the OA algorithm towards high-performing solutions.

2.2 Literature Review: Mixtures

Using a multicomponent mixture as an ORC working fluid is another route to improve ORC efficiency. For certain applications, such as low hot fluid temperatures and constrained cooling water cases, mixtures can lead to increased thermodynamic and economic performances (Chys et al., 2012; Lecompte et al., 2014; Oyewunmi et al., 2016). Particularly, the non-isothermal isobaric phase change for zeotropic mixtures (Figure 2), allows an improved thermal match with the heat source upon evaporation compared to pure working fluids with constant temperature profiles (Linke et al., 2015).

When considering mixtures, the lack of accurate property data models and the increased combinatorial complexity compared to pure fluids have led to optimisation approaches based on a pre-defined set of working fluid candidates (Andreasen et al., 2014; Molina-Thierry et al., 2015).

Papadopoulos et al. (2013) proposed the first and only CAMPD optimisation framework for binary mixtures where process conditions and working fluid structures of the two compounds are optimised simultaneously through simulated annealing (SA). Established GC methods (Joback and Reid, 1987) with semi-empirical mixing rules were used to estimate mixture properties.

While easily implementable, SA results in a distribution of nearly-optimal solutions and can suffer from long computational time (Pardalos et al., 2000). In addition, using (semi-)empirical mixing rules and interaction parameters for predicting thermophysical properties increases the uncertainty of the GC thermodynamic model (Linke et al., 2010). The lack of robust CAMPD frameworks for mixtures highlights the opportunity to apply the advanced methods developed for pure fluids and harness the advantages offered by mixtures as ORC working fluids, namely by using deterministic optimisation algorithms and state-of-the-art GC property prediction methods.

2.3 Objectives

The aim of this study is to work towards the synthesis of a complete CAMPD framework for ORCs which implements binary mixtures as working fluids and uses SAFT-γ Mie for accurate property prediction. An outer approximation algorithm would guarantee global optimality in determining the molecular structure of the components, mixture composition and ORC process conditions.

As an intermediate checkpoint to building this holistic framework, this work primarily addresses the ORC design primal problem, which determines optimal process conditions, given a choice of working fluid. Adaptations to the thermodynamic model are proposed to handle mixtures and composition is added as a decision variable in the optimisation.

Secondly, given the success of feasibility tests in pure fluid CAMPD optimisation (Bowskill et al. 2019), the focus shifts to modifying and building novel tests to aid the robustness of the future algorithm (Figure 1). Ultimately, every candidate binary mixture, consisting of two pure compounds generated by the master problem, would have to pass the feasibility tests for its performance to be considered in the primal problem.

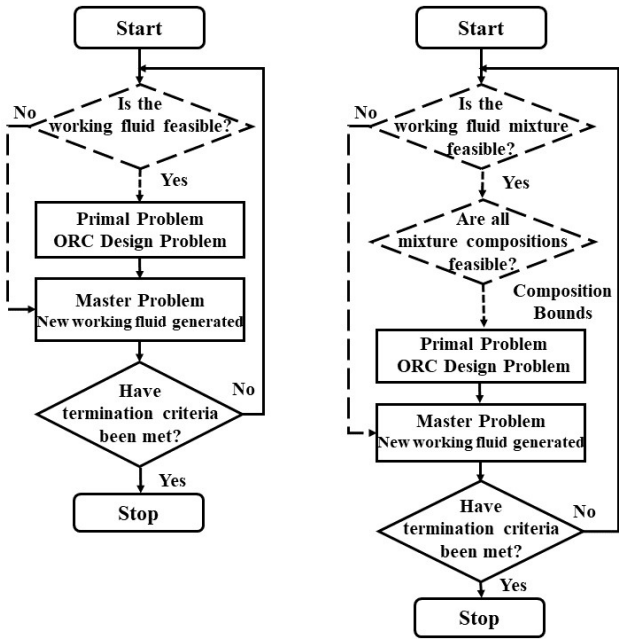


Figure 1: Outer Approximation (OA) Algorithm with feasibility tests proposed by Bowskill et al. for pure fluids (left) and augmented with a composition test for mixtures (right)

3 Methodology

3.1 Primal Problem Adaptation to Mixtures

The model of a simple ORC configuration (pump, evaporator, turbine and condenser), proposed by Bowskill et al. (2016) with a pure working fluid, is adapted to mixtures and implemented in gPROMS ModelBuilder. Figure 2 shows a temperature-entropy diagram for the process considered.

3.1.1 ORC process model

Isobaric evaporation and condensation are assumed. The cycle thus operates at a high and a low pressure regime defined by the evaporator pressure, P_{evap}, and condenser pressure, P_{cond}. The operating pressure values are constrained by absolute pressure bounds (P^L and P^U),

entered by the user, where P^L is set to 1 bar to avoid subatmospheric operation and P^U is 50 bar, due to high capital cost at high pressures. In addition, subcritical operation is ensured via Equations 1 and 2, where the lower and upper reduced pressure factors (P_r^L and P_r^U) are values between 0 and 1, set to 10^{-5} and 0.8 respectively.

$$\frac{P_{evap}}{P_{c,mix}} \leq P_r^U \tag{1}$$

$$\frac{P_{cond}}{P_{c,mix}} \geq P_r^L \tag{2}$$

As the gPROMS implementation of SAFT-γ Mie does not have a built-in method for critical pressure (P_c) evaluation, a standalone code developed by Bowskill et al. (2016) is used. However, this method cannot be integrated with the optimisation to model the mixture critical pressure as a function of composition z. Hence, $P_{c,mix}$ is approximated as a linear combination of the critical pressures of the pure components ($P_c(\mathbf{n_1})$, $P_c(\mathbf{n_2})$) entered as parameters.

$$P_{c,mix} = zP_c(\mathbf{n_1}) + (1-z)P_c(\mathbf{n_2}) \tag{3}$$

The turbine and pump operate adiabatically with a constant isentropic efficiency η . The quality of the vapor allowed at the turbine outlet is set to $f_{vap} = 1$ to avoid damage to the turbine.

The power output from each piece of equipment (pump and turbine) is evaluated as follows:

$$\dot{W} = \eta \frac{\dot{m}}{MW_{mix}}(h_{in} - h_{out}) \tag{4}$$

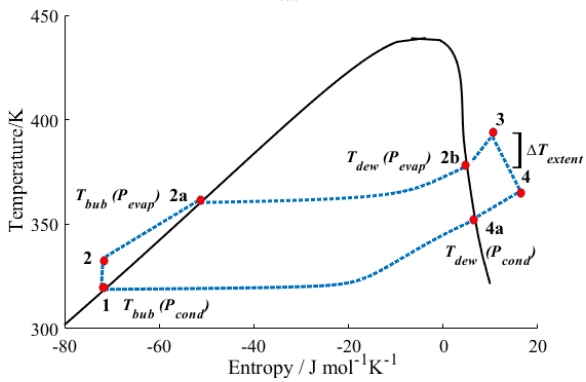


Figure 2: Temperature-entropy diagram of the ORC considered using a mixture as a working fluid. Superheat is allowed in the evaporator, no subcooling occurs at the condenser outlet.

The working fluid can be superheated in the evaporator by ΔT_{extent} (state 3, in Figure 2). Saturated liquid at bubble temperature is assumed at the condenser exit (state 1, in Figure 2). The heat source is modelled as a hot carrier fluid exchanging heat with the working fluid in a counter-current configuration . Constant heat flow capacity $\dot{M}c_p$ and inlet temperatures are defined for both the heat source and cooling water (summarised in Table 1). Heat exchange was modelled assuming no loss to the surrounding as shown by equations :

$$\dot{Q} = \frac{\dot{m}}{MW_{mix}}(h_{in} - h_{out}) \tag{5}$$

$$\dot{Q} = \dot{M}c_p(T_{in} - T_{out}) \tag{6}$$

To guarantee feasible heat transfer a minimum approach temperature ΔT_{min} of 10K is set in the heat exchangers via pinch point constraints. In the evaporator, these are given by comparing the heat source T_{hs} and working fluid T_{wf} temperatures , while the cooling water temperature T_{cw} and heat source temperature T_{hs} are compared in the condenser (Equation 7-8).

$$T_{hs} \geq T_{wf} + \Delta T_{min} \tag{7}$$

$$T_{cw} \leq T_{wf} + \Delta T_{min} \tag{8}$$

In the pure fluid model proposed by Bowskill et al. (2016), pinch point constraints were implemented at the inlet, outlet and bubble point in the evaporator (states 2,3 and 2a in Figure 2). A lower bound for the outlet temperature of the working fluid based on cooling water inlet temperature was the only constraint for the condenser.

For mixtures, additional constraints are required to account for the non-isothermal phase change behaviour. The non-violation of the minimum approach temperature is enforced at the inlet and outlet as well as the bubble and dew point of the working fluid for the evaporator (states 2, 2a, 2b, 3) and the condenser (states 4, 4a ,1). Interior point constraints within the vapor liquid equilibrium (VLE) region are also introduced (between states 2a-2b, and 4a-1). The number of pinch point constraints required to reliably reconstruct the temperature profile in the VLE region is determined in this work (Section 4.1).

3.1.2 Property Prediction Method

The state-of-the-art GC implementation of SAFT, SAFT-γ Mie, is used to determine the working fluid mixture properties (Papaioannou et al., 2014). These include: specific entropies, specific enthalpies, saturation temperatures, vapour fraction and composition of the liquid and vapour during phase change. The reliability of the SAFT-γ Mie framework to accurately describe VLE for binary mixtures over a wide range of conditions and systems, including ORC applications, has been demonstrated (Dufal et al., 2014; White et al., 2018).

The gPROMS implementation of SAFT-γ Mie breaks down around the critical point. Hence, the critical point of the pure components is indirectly computed with SAFT-γ Mie in a separate code. The critical pressure, P_c, and temperature, T_c, are found by increasing the pressure and comparing molar volumes of the gas, $v_{m,g}$, and liquid, $v_{m,l}$, phases. The pressures at which $v_{m,g} \approx v_{m,l}$ is recorded as P_c. The values obtained for the two pure compounds in the binary mixture ($P_c(\mathbf{n_1})$ and $P_c(\mathbf{n_2})$) are included in the primal problem as parameters.

Melting points were estimated using the GC method proposed by Hukkerikar et al. (2013).

3.2 Optimisation Formulation

$$\begin{aligned} &\min_{\mathbf{x}} f(\mathbf{x},\mathbf{n_1},\mathbf{n_2}) \\ &s.t. \mathbf{h}(\mathbf{x},\mathbf{n_1},\mathbf{n_2}) = \mathbf{0} \\ &\quad \mathbf{g}(\mathbf{x},\mathbf{n_1},\mathbf{n_2}) \leq \mathbf{0} \\ &\quad \mathbf{x} \in \mathbf{X} \\ &\quad \mathbf{n_1},\mathbf{n_2} \in \mathbf{N} \end{aligned} \tag{9}$$

The primal problem is a process optimisation problem over continuous variables $\mathbf{x}$ with fixed discrete variables $\mathbf{n_1},\mathbf{n_2}$. For the design of binary mixtures, the discrete variables correspond to the molecular structure of the two compounds in the mixture ($\mathbf{n_1}$ and $\mathbf{n_2}$).
The primal problem has five continuous degrees of freedom: condenser pressure (P_{cond}), evaporator pressure (P_{evap}), degree of superheat in the evaporator (ΔT_{extent}), working fluid mass flowrate ($\dot{m}$) and composition of component 1 in the binary mixture (z).
The net power output is used as the objective function to maximise the power extracted from the heat source.

$$\dot{W}_{net} = |\dot{W}_{turbine}| - |\dot{W}_{pump}| \tag{10}$$

The bounds on pressure from Test 0 and composition from Test 2 are used to narrow the search space for P_{cond}, P_{evap} and z. Assuming global optimality, the solution vector $\mathbf{x}$ presents the process conditions and mixture composition to reach the maximum power output for the set of compounds considered.

3.3 Feasibility Tests

Feasibility tests have been introduced in the OA framework prior to the primal problem to increase algorithm robustness. Test 0 and 1 proposed by Bowskill et. al (2016) for pure fluids are adapted to mixtures, to identify tighter bounds on the operating region and discard infeasible fluids combinations. A new test, Test 2, is proposed to narrow the bounds of feasible mixture compositions (Figure 1).

3.3.1 Test 0: Absolute Process Bounds

Process conditions are constrained by absolute upper and lower bounds on pressure (P^U, P^L) and temperature (T^U, T^L). In this work, P^L is set to 1 bar to avoid subatmospheric operation. P^U excludes operation above 50 bar due to cost considerations. T^U and T^L are determined by the heat source temperature and cooling water temperature.

In Test 0 the critical pressures of the two components selected to form a mixture, $P_C(\mathbf{n_1})$ and $P_C(\mathbf{n_2})$, are compared to the absolute pressure bounds. An updated set of pressure bounds $P^{LO}_{\mathbf{n_1},\mathbf{n_2}}$ and $P^{UO}_{\mathbf{n_1},\mathbf{n_2}}$ is obtained via the formulation:

$$P^{L0}_{\mathbf{n_1},\mathbf{n_2}} = max\{P^L, P^L_r min\{P_c(\mathbf{n_1}), P_c(\mathbf{n_2})\}\} \tag{11}$$

$$P^{U0}_{\mathbf{n_1},\mathbf{n_2}} = min\{P^U, P^U_r max\{P_c(\mathbf{n_1}), P_c(\mathbf{n_2})\}\} \tag{12}$$

Where P^L_r and P^U_r are minimum and maximum reduced pressure values. As the user-specified bound P^U usually allows pressure to rise above the critical pressure of the two compounds $\mathbf{n_1},\mathbf{n_2}$, this procedure in Test 0 narrows the bounds on pressure with respect to the pure component with the highest P_c.

The absolute temperature lower bound T^L, is compared to the highest melting point of the two compounds, $T_m(\mathbf{n_1})$ and $T_m(\mathbf{n_2})$. The lower bound T^L is updated to $T^{L0}_{\mathbf{n_1},\mathbf{n_2}}$ to avoid formation of a solid phase in the system:

$$T^{L0}_{\mathbf{n_1},\mathbf{n_2}} = max\{T^L, max\{T_m(\mathbf{n_1}), T_m(\mathbf{n_2})\} + \Delta T\} \tag{13}$$

After Test 0, the set of bounds on temperature and pressure is $P^{LO}_{\mathbf{n_1},\mathbf{n_2}}$, $P^{UO}_{\mathbf{n_1},\mathbf{n_2}}$, $T^{LO}_{\mathbf{n_1},\mathbf{n_2}}$, T^U, with the upper bound in temperature unconstrained by the test, given the possibility to superheat the working fluid in the evaporator. The feasibility of the new defined region is evaluated via the following two constraints:

$$T^{L0}_{\mathbf{n_1},\mathbf{n_2}} - T^U \leq 0 \tag{14}$$

$$P^{L0}_{\mathbf{n_1},\mathbf{n_2}} - P^{U0}_{\mathbf{n_1},\mathbf{n_2}} \leq 0 \tag{15}$$

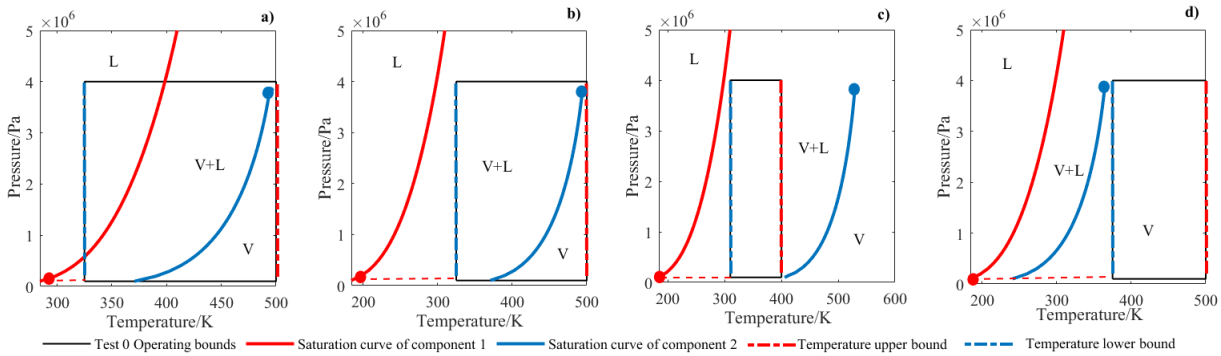


Figure 3: Generalised cases for Test 1. Cases (a)-(c) show feasible working fluid combinations given the overlap of the VLE region wih the operating bounds. This is verified in Test 1 with $T^{sat}_{min,\mathbf{n_1},\mathbf{n_2}}$(●) lower than the temperature upper bound (T^U) and $T^{sat}_{max,\mathbf{n_1},\mathbf{n_2}}$(●) higher than the temperature lower bound (T^{LO}) (Equations 18 and 19). In case (d) the operating bounds are restricted to the vapour phase region, making the fluid combination infeasible.

3.3.2 Test 1: Saturation Behaviour

The feasibility of phase change in the condenser and the evaporator is assessed in Test 1. As the saturation curves of a binary mixture lie between the saturation curves of the pure components, Test 1 is formulated by evaluating the saturation temperatures of the pure compounds.

The lowest saturation temperature at which a phase change can occur, denoted $T^{sat}_{min,\mathbf{n_1},\mathbf{n_2}}$, is effectively the lowest of the saturation temperatures of the two pure components, evaluated at the lower bound in pressure $P^{LO}_{\mathbf{n_1},\mathbf{n_2}}$. Similarly, the maximum saturation temperature for a vapour-liquid transition is $T^{sat}_{max,\mathbf{n_1},\mathbf{n_2}}$, which is computed at the upper bound P^{U0}, hence the formulations:

$$T^{sat}_{min,\mathbf{n_1},\mathbf{n_2}} = min\{T^{sat}(P^{L0}_{\mathbf{n_1},\mathbf{n_2}},\mathbf{n_1}), T^{sat}(P^{L0}_{\mathbf{n_1},\mathbf{n_2}},\mathbf{n_2})\} \quad (16)$$

$$T^{sat}_{max,\mathbf{n_1},\mathbf{n_2}} = max\{T^{sat}(P^{U0}_{\mathbf{n_1},\mathbf{n_2}},\mathbf{n_1}), T^{sat}(P^{U0}_{\mathbf{n_1},\mathbf{n_2}},\mathbf{n_2}),\} \quad (17)$$

If the minimum saturation temperature $T^{sat}_{min,\mathbf{n_1},\mathbf{n_2}}$ is higher than the temperature upper bound, T^{U0}, vapour cannot be formed and the compound combination $(\mathbf{n_1},\mathbf{n_2})$ is deemed infeasible as a working fluid mixture. Similarly, a combination of fluids with $T^{sat}_{max,\mathbf{n_1},\mathbf{n_2}}$ below the lower temperature bound, T^{L0}, implies that a vapour-liquid transition cannot occur in the condenser given the process bounds. Thus, for a mixture to pass the test the following two constraints must be satisfied:

$$T^{sat}_{min,\mathbf{n_1},\mathbf{n_2}} - T^{U} \leq 0 \quad (18)$$

$$T^{L0}_{\mathbf{n_1},\mathbf{n_2}} - T^{sat}_{max,\mathbf{n_1},\mathbf{n_2}} \leq 0 \quad (19)$$

Figure 3 illustrates different cases for Test 1 in a pressure-temperature diagram. For working fluid combinations in Figure 3 (a), (b) and (c), there is an overlap between the VLE region (delimited by the two saturation curves) and the box representing the process bounds from Test 0. These fluids thus pass Test 1. Figure 3 (d) shows a case where Equation 18 is violated, which implies that liquid cannot be formed for any sets of conditions within the process bounds. Analogously, vapour cannot be formed if Equation 19 is not satisfied. In such instances, Test 1 discards the candidate set of pure compounds as infeasible.

3.3.3 Test 2: Composition Test

For a binary mixture the dew and bubble saturation curves lie between the vapor pressure curves of the pure components. Thus, infeasible regions for the composition z arise when the two fluids considered are:

1. a feasible and an infeasible pure component (Figure 3 (b))
2. two infeasible fluids on either side of the process bounds (Figure 3 (c))

Test 2 identifies the feasible composition range given a set of two compounds.

For a binary mixture of composition z to be feasible as a working fluid, it is required that a liquid to vapour and vapour to liquid transition can occur given the process bounds. Graphically, this corresponds to both the bubble and dew curve intersecting the operating region defined by the process bounds ($P^{LO}_{\mathbf{n_1},\mathbf{n_2}}$, $P^{UO}_{\mathbf{n_1},\mathbf{n_2}}$, $T^{LO}_{\mathbf{n_1},\mathbf{n_2}}$, T^{U}) from Test 0.

Due to the monotonicity of the bubble and dew temperature curves with respect to pressure, Test 2 is formulated by evaluating the saturation temperature (T_{sat}) at the lower bound ($P^{LO}_{\mathbf{n_1},\mathbf{n_2}}$) and upper bound on pressure for a mixture with composition z, and comparing them to the process bounds on temperature. $P^{LO}_{\mathbf{n_1},\mathbf{n_2}}$ is defined by Test 0. As subcritical operation is assumed, the upper bound on pressure is updated to ($P^{U2}_{\mathbf{n_1},\mathbf{n_2},z}$) depending on the composition z as follows:

$$P^{U2}_{\mathbf{n_1},\mathbf{n_2},z} = P_{c,mix} P^{U}_{r} \quad (20)$$

Where $P_{c,mix}$ is the critical pressure of the mixture (Equation 3) and P^{U}_{r} is the upper reduced pressure bound set to 0.8.

Three cases are identified which would lead to the saturation curve intersecting the region defined by the process bounds (Figure 4).

$$T_{sat}(P^{U2}_{\mathbf{n_1},\mathbf{n_2},z},z) \geq T^{U} \;\wedge\; T_{sat}(P^{LO}_{\mathbf{n_1},\mathbf{n_2}},z) \leq T^{LO}_{\mathbf{n_1},\mathbf{n_2}} \quad (21)$$

$$T^{LO}_{\mathbf{n_1},\mathbf{n_2}} \leq T_{sat}(P^{U2}_{\mathbf{n_1},\mathbf{n_2},z},z) \leq T^{U} \quad (22)$$

$$T^{LO}_{\mathbf{n_1},\mathbf{n_2}} \leq T_{sat}(P^{LO}_{\mathbf{n_1},\mathbf{n_2}},z) \leq T^{U} \quad (23)$$

If any of the cases in Equations 21-23 is satisfied for both the dew and bubble temperature, the given mixture composition z is evaluated as feasible. This process is

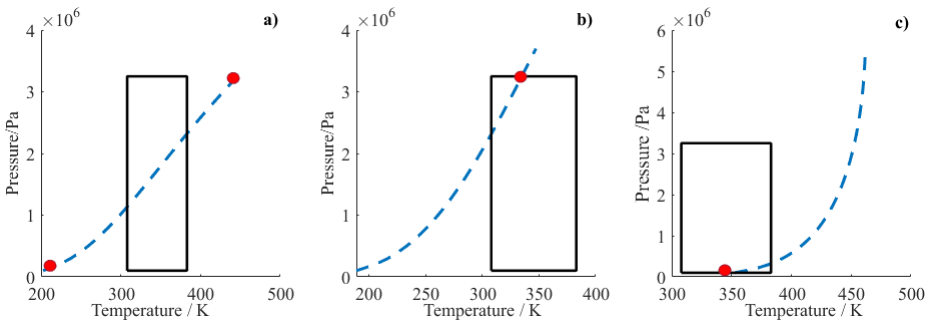


Figure 4: Generalised cases of saturation (dew or bubble) curves for binary mixtures with composition z (- - -). In cases (a)-(c) the T^{sat} evaluated at $P^{U2}_{\mathbf{n_1},\mathbf{n_2},z}$ and $P^{LO}_{\mathbf{n_1},\mathbf{n_2}}$ (●) satisfy the constraints in Equations 21-23 respectively. Test 2 identifies the intersection of the saturation curve with the operating region (in black). If case (a) (b) or (c) is met for both the dew and bubble curve that mixture composition z passes Test 2 as 'feasible'

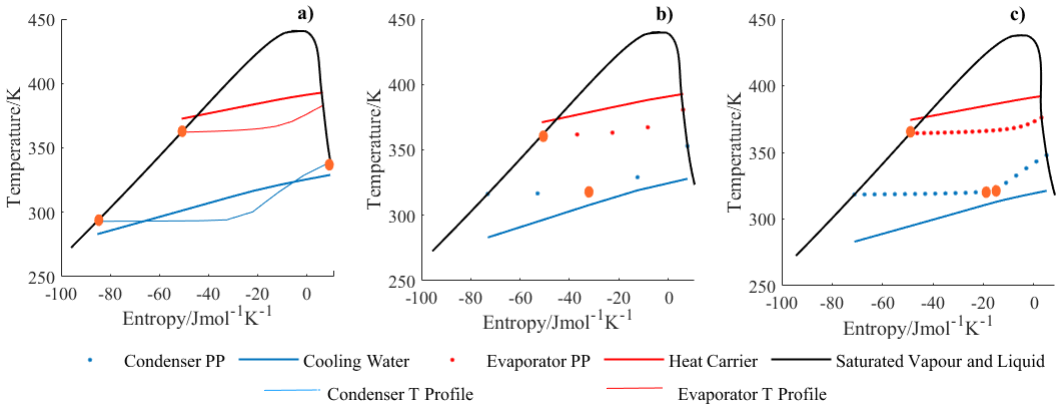


Figure 5: Change in the optimal ORC process conditions for butane-propanol as the number of pinch point (PP) constraints in the VLE region is varied (● evaporator PP, ● condenser PP) . Infeasible heat exchange occurs with no interior PP (a). The set of active constraints (●) changes when 5 (b) and 20 (c) PP computations are introduced in the VLE region.

repeated for the entire set of discrete compositions ($z \in [0, 1]$), tested in 0.1 increments in this work.

The solution from the composition test generates tighter bounds on the composition decision variable z. the initial guess on composition (z_0), set to 0.5 by default, is also updated to a feasible composition if required.

3.4 Process and Heat Source Specifications

Heat source specifications typical of geothermal applications are used in this study. A summary of the process specifications and constraints as used by Bowskill et al. (2016), with the addition of cooling water specifications to model heat exchange in the condenser is displayed in Table 1.

Table 1: Summary of ORC specifications used in this paper

Parameter	*Symbol*	*Units*	*Value*
Mass Flowrate (Heat Source)	$\dot{m}_{HS}$	kg s^{-1}	66
Heat Capacity (Heat Source)	$c_{p,HS}$	kJ kg$^{-1}K^{-1}$	4.2
Inlet Temperature (Heat Source)	$T_{hs,in}$	K	393
Mass Flowrate (Cooling Water)	$\dot{m}_{CW}$	kg s^{-1}	100
Heat Capacity (Cooling Water)	$c_{p,CW}$	kJ kg$^{-1}K^{-1}$	4.19
Inlet Temperature (Cooling Water)	$T_{hs,in}$	K	283
Min. Approach Temperature	$_{min}$	K	10
Min. Abs. Temperature	T^L	K	293
Max. Abs. Temperature	T^U	K	383
Min. Abs. Pressure	P^L	bar	1
Max. Abs. Pressure	P^U	bar	50
Min. Red. Pressure	P_r^L	-	10^{-5}
Max. Red. Pressure	P_r^U	-	0.8
Isentropic Efficiency (turbine)	η_t	-	0.8
Isentropic Efficiency (pump)	η_p	-	0.9

As the master problem is not implemented in this work, the discrete variables defining the molecular structures of the two compounds in the binary mixture ($\mathbf{n_1}$ and $\mathbf{n_2}$) are entered by the user.

4 Results and Discussion

4.1 Heat Exchanger Modeling

As discussed in Section 3.1.1, pinch point constraints (Equations 7-8) are added in the VLE region to ensure feasible heat transfer along the heat exchangers. The number of interior point constraints required is determined in this section.

When pinch point constraints are only enforced at the inlet/outlet of the unit, and bubble/dew points of the working fluid, the optimal process conditions found by the primal problem can lead to violation of the minimum approach temperature (ΔT_{min}) in the VLE. As shown in Figure 5 (a), when interior point constraints are not implemented in the VLE region the primal problem output for a butane-propanol mixture results in infeasible heat transfer in the condenser.

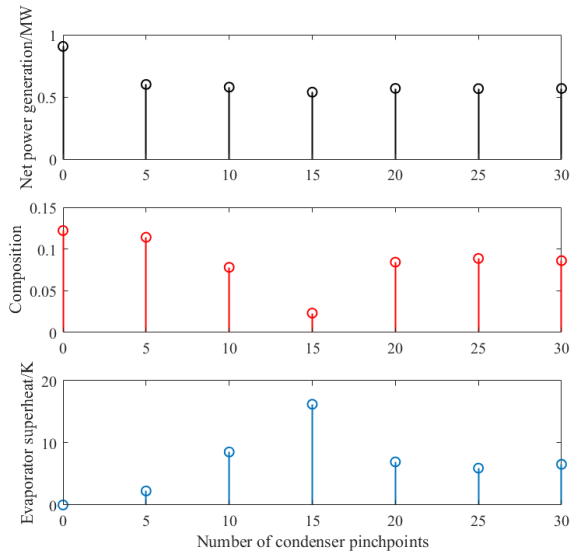


Figure 6: Sensitivity of the optimisation output (Net Power Generation (MW), Mixture composition z, Evaporator Superheat ΔT_{extent} (K)) to the number of interior pinch point constraints in the VLE region during heat exchange. The optimum converges after 20 points are introduced.

A sensitivity analysis is conducted to assess the optimal number of interior constraints required. High variability in the primal problem output, in terms of optimal composition

and evaporator superheat is observed with 5 to 15 constraints, due to the insufficient approximation of the condensation profile in the VLE. The results are explained by the activation of a different set of constraints at the solution point for the three cases considered (Figure 5).

As shown in Figure 6, beyond 20 interior points, the optimisation output converges. Figure 5 (c) illustrates the reliable reconstruction of the non-isothermal isobar achieved for the butane-propanol mixture with 20 interior points. 20 pinch point constraints are thus implemented in the VLE region.

4.2 Impact of Tests Implementation

The impact of implementing the tests developed in Section 3.3 is illustrated through the butane-decane binary system. Butane is feasible as a pure working fluid while decane is infeasible (Figure 7) given the process conditions utilised (Table 1).

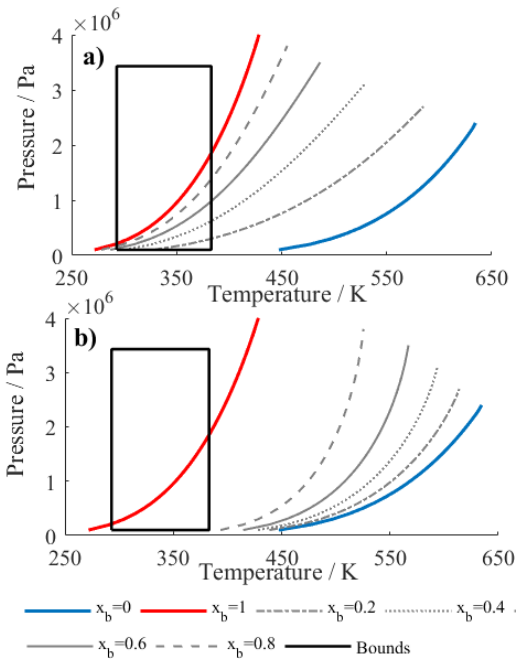


Figure 7: Bubble (a) and dew (b) curves for mixtures of butane (red) and decane (blue) with composition of butane (x_b) in the range $[0,1]$. Overlap of the bubble and dew curves with the process bound (box) indicates the feasibility of the given composition .

Test 0 reduces the maximum pressure bound $P^{U0}_{\mathbf{n1,n2}}$ from 50 bar to 34.3 bar (the reduced critical pressure of butane). Test 1 verifies the possibility for VLE given the pressure and temperature bounds. Finally Test 2 generates tighter bounds, $z \in [0.8-1.0]$, on feasible butane-decane compositions. The initial guess for the composition decision variable (z_0) is also updated from the default value (0.5) to 0.9. The impact of implementing the results from tests 0,1 and 2 on the optimisation output is summarised in Table 2.

The optimisation fails when no tests are implemented. Tightening the pressure bounds through Tests 0-1 leads to a trivial solution with a net power output of $2.8\,10^{-14} MW$. Updating the composition bounds and z_0 with the results from Test 2 leads to an optimal solution with a meaningful power output. Implementing the results from Tests 0-2 simultaneously results in the same optimal solution but lowers the computation time from 220s to 116s. In this case, finding tighter bounds on composition with Test 2 is key to converging to the global optimum.

Table 2: Impact of feasibility tests (Test 0/Test1) and composition tests (Test 2) on the optimisation output for a butane-decane binary system

Test Implemented	**None**	**0 – 1**	**2**	**0 – 1 – 2**
Status	Fail	Optimal	Optimal	Optimal
Net Power Output $\dot{W}_{net}(MW)$	-	$2.8\,10^{-14}$	0.56	0.56
CPU time (s)	-	12	220	116

This example illustrates the success of the feasibility and composition tests in enhancing algorithm robustness. The adaptation of the primal problem to mixtures increases the level of complexity in the gPROMS ORC model. However, by eliminating large regions of infeasibility from the search space, the introduction of tests to the OA framework allows the initialisation of the highly non-linear primal problem from a starting point which is closer to the optimal solution thus aiding with convergence.

4.3 Feasible-Feasible Binary Mixtures

The top 10 performing pure working fluids for the case study considered in this paper were reported by Bowskill et al. (2016) through implementation of the full CAMPD algorithm . Using these fluids as input to the primal problem developed in this work leads to a slight change in the hierarchy as pinch point constraints are not included in the inlet of the condenser in the model proposed by Bowskill et al. (2016). The performance of 5 arbitrary combinations of the top 1-5 and 10 pure working fluids was assessed by optimising the mixture composition and ORC process conditions in the primal problem. Table 3 summarises the net power output ($\dot{W}_{net}$) obtained.

Some mixtures outperform the pure fluids: the optimal composition ($z_{propane} = 0.61$) for a propane and methyl ethyl ether (MEE) mixture delivers an 18% increase in $\dot{W}_{net}$ compared to pure propane. Remarkably, when the top ranked working fluid, propene, is mixed with pent-1-ene, ranked 10, the $\dot{W}_{net}$ is higher by 33% compared to pure propene. Conversely, despite the promising performance of pure butane and MEE, the objective function is reduced when combined as a mixture.

It is thus challenging to extract a trend predicting performance of mixtures based solely on that of the individual fluids employed. Hence, the need to implement a full CAMPD algorithm to determine the molecular structure of the fluids included in the mixture is underlined.

Table 3: Comparison of the performance of top pure working fluids identified by Bowskill et al. (2016) and arbitrary binary combinations in mixtures, with optimised composition and process conditions.

Rank	Top Pure Fluids		Top Mixtures		
	Name	$\dot{W}_{net}(MW)$	Name	$\dot{W}_{net}(MW)$	Optimal z
1	Propene	0.900	Propene/Pent-1-ene	1.196	0.70
2	Propane	0.885	Propane-Methyl Ethyl Ether	1.076	0.61
3	But-1-ene	0.856	Butane-Propane	1.015	0.27
4	Butane	0.853	But-1-ene-Methyl Ethyl Ether	0.866	0.58
5	Methyl Ethyl Ether	0.852	Butane-Methyl Ethyl Ether	0.838	0.24
10	Pent-1-ene	0.846			

4.4 Infeasible-Feasible Binary Mixtures

The impact of mixing a compound which would be infeasible as a pure working fluid with one that is feasible (here classified as infeasible-feasible mixtures) is explored.

4.4.1 Case study: Butane - Octane/Nonane/Decane

Three binary systems composed of a feasible working fluid (butane) and an infeasible working fluid (octane/nonane/decane) are considered (Figure 8).

The optimal composition of the butane-octane, butane-nonane and butane-decane binary systems that maximises the net power output ($\dot{W}_{net}$) is obtained with the tests and primal problem formulation adapted for mixtures, where composition of the mixture (z) and process conditions are optimised.

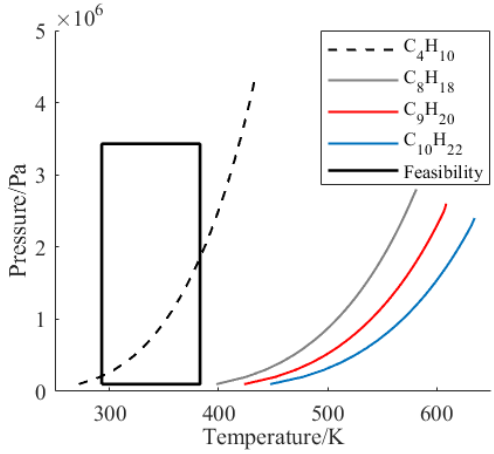


Figure 8: Saturation curves of pure butane, octane, nonane and decane. The absence of overlap with the feasible P-T region for octane, nonane and decane shows that they are infeasible as pure working fluids.

Two cases are considered, in which the cooling water flowrate ($\dot{m}_{CW}$) in the condenser is:

(i) fixed at $\dot{m}_{CW} = 40\,kg\,s^{-1}$

(ii) fixed at $\dot{m}_{CW} = 400\,kg\,s^{-1}$

The power output for the mixtures are compared with the maximum power output obtained when optimising the process conditions for pure butane as a working fluid. Results are summarised in Figure 9.

4.4.2 Effect on the net power output $\dot{W}_{net}$

In both cases the introduction of an infeasible working fluid increases the maximum net power output compared to the pure fluid at optimal process conditions. It is apparent that the closer the infeasible fluid is to the feasible region the higher the increase in $\dot{W}_{net}$ compared to pure butane. For case (i) 3mol% octane content in the mixture leads to a 22.3% increase in net power output, whilst a 10.1% increase is achieved for a 1mol% decane-butane mixture.

Another point of interest is the low concentration of the infeasible fluid in the optimal composition. The primal problem implemented does not include a costing model. Hence, the impact of potential liquid-liquid equilibrium on equipment costs, and the possible increased heat exchanger costs for zeotropic mixtures reported in literature is not considered (Oyewunmi and Markides, 2016). Given the likely increase in capital cost linked to the use of mixtures, a constraint could be implemented to restrict the minimum composition of a component in the final mixture to 5-10 mol%.

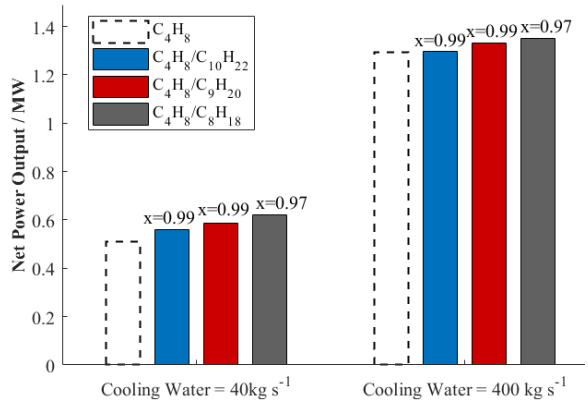


Figure 9: Net power output (MW) for pure butane, and mixtures of butane with octane/nonane/decane. Case (i) $\dot{m}_{CW} = 40kgs^{-1}$ Case (ii) $\dot{m}_{CW} = 400kgs^{-1}$. The increase in power output for mixtures is greater in case (i)

4.4.3 Impact of cooling water availability

From the comparison of case (i) and (ii), binary mixtures are particularly attractive alternatives to pure fluids when the availability of cooling water is restricted.

In case (ii) the increase in net power output compared to pure butane is +4.5% for butane-octane and +0.39% for a butane-decane mixture. The performance enhancement is less significant than case (i) where +22.3% and +10.1% were achieved for butane-octane and butane-decane respectively. This is in line with the thermoeconomic assessment of White et al. (2018) which found that mixtures only outperformed pure fluids as ORC working fluids in

cases where cooling water was constrained.

This case study highlights the possibility for infeasible pure substances to enhance the net power output when mixed with feasible pure substances, especially when cooling water availability is limited. This supports the importance of not discarding infeasible pure working fluids as candidates for binary mixtures when implementing the full CAMPD algorithm.

4.5 Algorithm Limitations

4.5.1 Limitation of P_c estimation

When considering feasible-infeasible (Figure 3(b)) or infeasible-infeasible mixtures (Figure 3(c)) Test 2 fails to complete due to numerical issues in some cases. Reducing the value of the reduced maximum pressure P_r^U allows Test 2 to complete successfully. Numerical failures are not encountered for feasible-feasible mixtures.

Out of 15 arbitrary feasible-infeasible and infeasible-infeasible mixtures of alkanes and alkenes (C_2 to C_{10}), 6 mixtures passed Test 2 using the default P_r^U values of 0.8. For the remaining mixtures P_r^U value was lowered by -0.5 increments and the P_r^U value at which the test completed was recorded (Table 4). Only infeasible-infeasible mixtures (such as ethane-nonane) required the P_r^U value to be lowered below 0.7.

Table 4: Maximum value of P_r^U (Reduced Critical Pressure) at which Test 2 succeeds without numerical issues for 15 binary systems (9 infeasible-feasible mixtures and 6 infeasible-infeasible mixtures)

P_r^U	**0.6**	**0.65**	**0.7**	**0.75**	**8**
Number of Mixtures	1	4	1	1	6

The gPROMS output mentions the failure of gSAFT (the gPROMS implementation of SAFT-γ Mie) as the cause of the test non-success. In Test 2 (Section 3.3.3) the saturation temperature of a mixture is evaluated at $P^{U2}_{\mathbf{n_1},\mathbf{n_2},z}$ computed for each composition via Equation 20. $P^{U2}_{\mathbf{n_1},\mathbf{n_2},z}$ depends on the critical pressure of the mixture ($P_{c,mix}$) and P_r^U. As gSAFT breaks down past the critical point it is likely that the gSAFT failures are caused by an overestimate of the critical pressure $P_{c,mix}$. This reflects the limitation of the linear approximation used for $P_{c,mix}$ (Equation 3.) and motivates the need to develop methods to accurately predict $P_{c,mix}$ in gPROMS which can be integrated in the optimisation.

4.5.2 Composition Initial Guess z_0 for mixtures

The composition initial guess z_0 is a key parameter when solving the optimisation due to the highly non-linear nature of the formulated primal problem.

The butane-propanol mixture is an example of a feasible-feasible mixture where the entire composition range ($z \in [0-1]$) corresponds to viable working mixtures in the process. As shown by Table 5, the optimisation converges to the global solution point for most initial guesses z_0. The solver failure for a z_0 of 0.4 could be caused by a point of discontinuity in the objective function with respect to the variable z. The non-convexity of the formulation is also apparent with the detection of a local optimum for a z_0 of 0.9, with a significantly lower net power output.

For feasible-infeasible mixtures, the completion of the optimisation is also impacted by the composition initial guess. Despite initialising it within the feasible z range obtained from Test 2, certain values of z_0 lead to failed optimisations. As shown in Table 5, starting from a z_0 that is closer to the actual optimal value of z can lead to optimisation failure.

Table 5: Optimisation status for different initial composition guesses (z_0) within the feasible butane composition range for two feasible-infeasible mixtures (butane-octane, butane-decane) and a feasible-feasible mixture (butane-propanol)

	Butane-Octane		**Butane-Decane**		**Butane-Propanol**	
Feasible z_b	[0.3,1]		[0.8,1]		[0,1]	
Optimal z_b	0.95		0.99		0.10	
	z_0	Status	z_0	Status	z_0	Status
Output	0.7	Fail	0.9	Optimal	0.1-0.3	Optimal
	0.8	Optimal	0.95	Fail	0.4	Fail
	0.9	Fail	0.99	Fail	0.5-0.8	Optimal
	1.0	Fail	1.0	Fail	0.9	Local

When updating z_0 in Test 2, the mid-range composition in the feasible z interval is chosen. If this value leads to optimisation failure due to numerical issues, a new composition within the range is tested as the initial guess. Such procedures for composition initialisation should be formalised and automated if the entire CAMPD OA algorithm is implemented.

5 Conclusion and Outlook

5.1 Conclusion

In this work, the primal problem of a CAMPD framework is adapted to binary mixtures as working fluids to identify optimal mixture composition and process conditions to achieve maximum net power output. SAFT-γ Mie is used for accurate mixture property prediction and feasibility tests are introduced to the traditional OA algorithm.

The adaptation and introduction of the feasibility tests preceding the primal problem increase the future algorithm robustness and reduce the overall computational time. By eliminating infeasible binary combinations of fluids (Tests 0-1) and infeasible regions of the composition search space (Test 2), feasibility tests allow the initialisation of the NLP from a point closer to the optimal solution thus aiding the convergence of the primal problem.

Some mixtures lead to significant improvement of the objective function compared to pure compounds. However, the unpredictability of mixture performance based on the pure fluids demonstrates the necessity of a full CAMPD

algorithm to identify promising mixtures.

The potential for mixtures of a feasible pure fluid and infeasible pure fluid to enhance the thermodynamic performance of the ORC is illustrated through the butane-octane/nonane/decane case study. Hence, when implementing the full CAMPD alogrithm, infeasible pure working fluids should not be discarded as potential candidates for mixtures.

5.2 Future Work

As discussed in Section 4.5, developing an integrated method to estimate the critical pressure for mixtures ($P_{c,mix}$) in gSAFT is key as $P_{c,mix}$ determines the upper bound on pressure in the primal problem. While the tests increase algorithm robustness, the initialisation procedure needs to be automated for the composition decision variable before implementing the full OA algorithm.

A cost model could also be included in the ORC model. Although this work highlighted the potential for zeotropic mixtures to increase ORCs thermodynamic performance, they could also lead to increased capital cost compared to pure working fluids. In particular, the impact of the temperature glide on heat exchanger costs and potential liquid-liquid equilibrium on equipment cost should be investigated. The use of azeotropic mixtures could also be studied as the presence of azeotropes in heat exchangers has been linked with higher thermodynamic efficiency and lower capital costs compared to zeotropic mixtures (Oyewunmi et al., 2016).

Finally, in Section 4.4.3, mixtures are found to be particularly attractive when cooling water availability is limited. Once a costing model is included, further work could be done to identify cases (heat source types, cooling water availability or cycle topology) where using mixtures would bring significant thermoeconomic benefits.

6 Acknowledgements

We want to thank Lauren Lee and David Bowskill for their availability, patience and guidance throughout the entirety of the project.

References

Andreasen, J., Larsen, U., Knudsen, T., Pierobon, L. and Haglind, F. (2014). Selection and optimization of pure and mixed working fluids for low grade heat utilization using organic Rankine cycles. *Energy*, 73, pp.204-213.

Bowskill, D., Tropp, E., Gopinath, S., Jackson, G., Galindo, A. and Adjiman, C. (Accepted). Beyond a heuristic analysis: Integration of process and working fluid design for organic Rankine cycle.

Dufal, S., Papaioannou, V., Sadeqzadeh, M., Pogiatzis, T., Chremos, A., Adjiman, C., Jackson, G. and Galindo, A. (2014). Prediction of Thermodynamic Properties and Phase Behavior of Fluids and Mixtures with the SAFT- Mie Group-Contribution Equation of State. Journal of Chemical Engineering Data, 59(10), pp.3272-3288.

Hukkerikar, A.S., Sin, G., Abildskov, J., Sarup, B. and Gani, R. (2013). Development of pure component property models for chemical product-process design and analysis. Technical University of Denmark, Department of Chemical and Biochemical Engineering, Kgs. Lyngby.

Joback, K. and Reid, R. (1987). Estimation of Pure Component Properties from Group-Contributions. *Chemical Engineering Communications*, 57(1-6), pp.233-243.

Lecompte, S., Ameel, B., Ziviani, D., van den Broek, M. and De Paepe, M. (2014). Exergy analysis of zeotropic mixtures as working fluids in Organic Rankine Cycles. *Energy Conversion and Management*, 85, pp.727-739.

Linke, P., Papadopoulos, A. and Seferlis, P. (2015). Systematic Methods for Working Fluid Selection and the Design, Integration and Control of Organic Rankine Cycles—A Review. *Energies*, 8(6), pp.4755-4801.

Macchi, E., Astolfi, M.(2017). Organic Rankine Cycle (ORC) Power Systems. *Technologies and Applications*, pp. 3-24.

Markides, C.(2015). Low-Concentration Solar-Power Systems Based on Organic Rankine Cycles for Distributed-Scale Applications: Overview and Further Developments. *Frontiers in Energy Research*, 3(47), pp. 1-16.

Molina-Thierry, D. and Flores-Tlacuahuac, A. (2015). Simultaneous Optimal Design of Organic Mixtures and Rankine Cycles for Low-Temperature Energy Recovery. *Industrial Engineering Chemistry Research*, 54(13), pp.3367-3383.

Oyewunmi, O., Taleb, A., Haslam, A. and Markides, C. (2016). On the use of SAFT-VR Mie for assessing large-glide fluorocarbon working-fluid mixtures in organic Rankine cycles. *Applied Energy*, 163, pp.263-282.M.

Oyewunmi, O. and Markides, C. (2016). Thermo-Economic and Heat Transfer Optimization of Working-Fluid Mixtures in a Low-Temperature Organic Rankine Cycle System. *Energies*, 9(6), p.448.

Oyewunmi, O., Kirmse, C. and Markides, C. (2016). Heat Exchanger Analysis of Azeotropes Heat Exchanger Mixtures in Organic Rankine Cycles, *UK Heat Transfer Conference*

Palma-Flores, O., Flores-Tlacuahuac, A. and Canseco-Melchor, G. (2015). Optimal molecular design of working fluids for sustainable low-temperature energy recovery.*Computers & Chemical Engineering*, 72, pp.334-349.

Papadopoulos, A., Stijepovic, M. and Linke, P. (2010). On the systematic design and selection of optimal working fluids for Organic Rankine Cycles. *Applied Thermal Engineering*, 30(6-7), pp.760-769.

Papadopoulos, A., Stijepovic, M., Linke, P., Seferlis, P. and Voutetakis, S. (2013). Toward Optimum Working Fluid Mixtures for Organic Rankine Cycles using Molecular Design and Sensitivity Analysis. Industrial Engineering Chemistry Research, 52(34), pp.12116-12133.

Papadopoulos, A., Tsivintzelis, I., Linke, P. and Seferlis, P. (2018). Computer-Aided Molecular Design: Fundamentals, Methods, and Applications. *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*.

Papaioannou, V., Lafitte, T., Avendaño, C., Adjiman, C., Jackson, G., Müller, E. and Galindo, A. (2014). Group contribution methodology based on the statistical associating fluid theory for heteronuclear molecules formed from Mie segments. The Journal of Chemical Physics, 140(5), p.054107.

Pardalos, P., Romeijn, H. and Tuy, H. (2000). Recent developments and trends in global optimization. *Journal of Computational and Applied Mathematics*, 124(1-2), pp.209-228.

Schilling, J., Lampe, M., Gross, J., Bardow, A. (2017). 1-stage CoMT-CAMD: An approach for integrated design of ORC process and wrokign fluid using PC-SAFT, *Chemical Engineering Science*, 159, pp.217-230.

White, M., Oyewunmi, O., Chatzopoulou, M., Pantaleo, A., Haslam, A. and Markides, C. (2018). Computer-aided working-fluid design, thermodynamic optimisation and thermoeconomic assessment of ORC systems for waste-heat recovery. Energy, 161, pp.1181-1198.

Zhang, X., Wu, L., Wang, X., Ju, G. (2016). Comparative study of waste heat steam SRC, ORC and S-ORC power generation systems in medium-low temperature. *Applied Thermal Engineering*, 106, pp.1427-1439

Evaluation of turbidity and pH technologies for Multi-Functional Micro Sensors (MFMS)

Ane García Gauna and Sofía Mediavilla Madrigal

Department of Chemical Engineering, Imperial College London, U.K.

Company confidential

Abstract

The transmitter industry has been stagnant for the last few decades, where little significant innovations have occurred. Multi-Functional Micro Sensors (MFMS) intend to revolutionise this industry by implementing a technology capable of making wireless measurements of several properties within a compact size. The work presented aims to assess different commercial pH sensors as well as an in-house built multi-beam turbidity sensor to determine whether they can compete with current existing products used in drinking-water treatment plants. It evaluates the possible implementation of the most suitable technologies to measure pH and turbidity in the next MFMS prototype, which is being developed in collaboration with a major instrumentation and control company. These devices were tested on a number of solutions and the results obtained were compared with a reference meter. The recommended pH sensor was the PyroScience Optical pH miniprobe due to its accuracy and fast response time to reach a stable measurement. The turbidity prototype was able to measure high values of turbidity (greater than 40 NTU). However, no clear trend could be observed for the range of interest (0-10 NTU), thus further testing and sensor design modifications are required.

1 Introduction

There have been few major developments in the design of modern transmitters over the last several decades. Besides from being large in size and expensive to maintain, they are cable dependent and only able to measure one property. In an attempt to remodel this industry, Multi-Functional Micro-Sensors (MFMS) for process monitoring are being developed. MFMS aim to make simultaneous wireless measurements of several properties within a small size, whilst being inexpensive to manufacture, accurate and reliable.

The overall purpose of the project, which is in collaboration between several Imperial College London departments and a major control and instrumentation company, is to design a MFMS for drinking-water applications that can compete with the existing technologies. A prototype that fulfills the overall aim is currently available and is able to measure four properties (Williams, 2018). The main focus of our research was to determine the best pH and turbidity technologies to incorporate in the following MFMS prototype by evaluating several commercial pH sensors together with an in-house assembled multi-beam turbidity sensor prototype and comparing the results obtained with a reference meter.

pH is an inherent property of aqueous solutions that measures the concentration of hydrogen ions, determining whether it is acidic, neutral or alkaline (*pH and Water*). Turbidity describes the haziness of a fluid caused by the number of suspended particles and other light-absorbing materials present. When the transmission of light in water is inhibited, this water sample is considered to be turbid (Kitchener et al., 2017).

Nephelometry and turbidimetry techniques were used to quantify the turbidity of a solution, the former measuring side-scattering and the latter transmission.

According to drinking-water standards, the acceptable pH value must be between 6.5 and 9.5 and turbidity must be lower than 4 Nephelometric Turbidity Units (NTU) (United Utilities, 2019; Johnson, 1961).

2 Materials

2.1 pH

2.1.1 Chemicals

In-house prepared buffer solutions
Buffer solutions were prepared by in-house mixing the following chemicals:

Potassium phosphate monobasic (KH_2PO_4) – analytical reagent of molar mass 136.09 g/mol and purity $>99\%$ from Sigma-Aldrich (St. Louis, MO, USA).

Sodium Hydroxide (NaOH) – analytical reagent of molar mass 40g/mol and purity 98.7% from Thermo Fisher Scientific (Loughborough, Leicestershire, UK).

Deionised (DI) water – DI water from a tap in the lab was filtered by a Millipore Stericup Quick Release Filter with a polyethersulfone membrane (PES) of $0.22\,\mu\mathrm{m}$ pore size.

Certified pH Buffer Solution
Potassium-based certified pH Buffer Solutions supplied by Paragon Scientific Ltd. (Prenton, Wirral, UK) with a pH value of 4, 7 and 10 with Fungicide, coloured and accurate within $\pm$ 0.01 at 25 °C were tested (*Buffer Solution, pH 10.00, Color-Coded Blue (Certified)*; *Buffer Solution, pH 7.00, Color-Coded Yellow (Certified)*; *Buffer Solution, pH 4.00, Color-Coded Red (Certified)*). Analysis with these solutions can be traced to recognised national standards and to units of measurement realised at the National Physical Laboratory (NPL) and National Institute of Standards and Technology (NIST) (Paragon Scientific Ltd, 2017).

2.1.2 pH Instruments

Mettler Toledo FE20 FiveEasy Benchtop pH Meter

This was used as the reference sensor. It ranges from pH 0 to 14. The temperature was manually set to 20 °C.

The measurements are based on the reaction between the glass membrane located in the probe, which is sensitive to hydrogen ions and the tested solution. The pH-sensitive electrode response depends on the hydrogen ion concentration and hence gives a signal that is determined by how acidic/alkaline the solution is. The observed potential of the electrode is compared to a reference one, which is not responsive to hydrogen ion concentration changes thus will always deliver a constant potential. The difference in potential is then utilised to calculate the pH of the analyte. Both electrodes are currently combined in one in which the pH-sensitive glass electrode is concentrically surrounded by the reference electrode filled with a reference electrolyte (Mettler Toledo, 2013).

PyroScience pH Sensor

This is an optical sensor whose pH range is 6 to 8. It is only 3 mm in diameter and it has an optical isolation protecting the pH measurement from disturbances at the sensing tip. The sensor is designed to minimise the cross-sensitivity effects due to the ionic strength of the solutions and is able to measure for long periods of time due to its considerable photo-stability (PyroScience GmbH, 2019).

Its measuring principle is based on REDFLASH technology which employs red light to excite REDFLASH indicators located in a fluorescence dye at the tip of the sensor which will show oxygen-dependent luminescence in the near infrared (NIR) spectrum region. The optical response will decrease with increasing oxygen molecules bonding to the indicators due to the quenching effect. The excited light as well as the NIR emission is transmitted through the optical fiber located inside the whole length of the sensor (*Piccolo2 Fiber-Optic Oxygen Meter User Manual*). The dye is immobilised in a polymer matrix and applied to a carrier material to be used as sensor spots. The red light is chosen because it can significantly reduce interferences caused by autofluorescence and stress in biological systems as compared with blue light (Idronaut S.r.l, 2019).

The change in optical properties as shown by the NIR emission is then measured by Piccolo2-OEM, which is an optical oxygen meter that includes a temperature port where a temperature sensor can be connected and used to automatically compensate for temperature changes. This is finally connected to a computer. The calibration and testing of the module is performed simultaneously through two softwares developed by the same company, Pyro Oxygen Logger and FireSting pH Logger (*Piccolo2-OEM Fiber-Optic Oxygen Meter OEM-Module with UART-Interface Manual*; *pH Sensor Manual - Piccolo2-OEM*).

Aquaread AP-2000

A commercial Aquaread Ltd. sensor which can measure both pH and turbidity was also utilised. It has the largest pH range (pH 0-14) out of the sensors tested. However, due to its large turbidity range (0-3000 NTU) and inaccuracy measuring low values, it was decided to only run pH-based experiments with it.

This multi-parameter water monitoring probe comes with several sensors and can be customised to add more to it. The probe has a built-in combined pH and Oxidation-Reduction Potential (ORP) sensor, a electrical conductivity (EC) meter followed by an optical dissolved oxygen (DO) cap and two auxiliary sensor ports, an Ion Selective Electrode (ISE) socket and an optical electrode socket. The latter can accommodate a turbidity meter.

The measuring principle of the combined pH and ORP sensor is as follows: the pH sensor measures the difference in electrical potential between two electrodes, a pH and a reference one, whereas the ORP sensor detects the changes in REDOX potential of the ORP electrode, which is based on the tendency of a chemical substance to oxidise/reduce another, and compares it with the reference electrode stable signal. Both sensors share a common reference electrode. The principle of this combined probe is similar to the measuring principle of Mettler Toledo pH meter (Peacock, 2014). The sensor was calibrated and data was logged from the probe on a PC by a testing software developed by the same company called AquaCal (*AquaCal – Aquaprobe Live Data & Calibration Utility with Report Generation*).

Microsens MSFET 3332

This sensor consists of a pH-Ion Selective Field-Effect Transistor (pH-ISFET) sensing component, an integrated miniature solid-state Ag/AgCl reference electrode and an electronic circuit assembled onto a miniature printed circuit board (PCB). More details about the electronics can be found in Microsens SA, 2017. Its typical pH testing range is 2 to 11.

The field-effect transistor is used to measure the ion concentration in a solution. The integrated circuitry runs at a constant current/voltage through the sensor. When the hydrogen ion concentration changes, the current through the transistor changes accordingly and thus an output signal proportional to the pH value of the tested analyte is obtained (*Glass pH electrode vs ISFET pH Electrode*).

In order to calibrate and log the data, a Laboratory Virtual Instrument Engineering Workbench (LabVIEW) program was developed to collect the voltage signals from the transistor. LabVIEW is commonly used for applications that require testing, measuring, and control of process parameters with rapid access to hardware and data insights (*LabVIEW 2019*).

Table 1: pH sensor model summary

Sensor Model	Type	Range
Mettler Toledo	Combined Electrodes	0-14
PyroScience	Optical	6-8
Aquaprobe	Combined pH/ORP	0-14
Microsens	ISFET	2-11

2.2 Turbidity

Currently, there are many turbidity sensors that claim they are able to operate over a wide range. However, there has been an ongoing challenge in the industry for several decades to accurately measure low values of turbidity. In an attempt to overcome this, a turbidity prototype sensor to measure low ranges was assembled by Dr Chao Yuan and collaborators, which was subsequently tested against an ABB Aztec ATS430 Sensor. The aim was to determine whether the prototype sensor is competitive with the current commercial sensors and to assess if its technology can be implemented in the following MFMS prototype.

2.2.1 Chemicals

Turbidity 1000 NTU Calibration Standard
Formazin turbidity standard is accepted as the primary method to calibrate turbidity sensors for the chemical examination of water and wastes. Formazin is produced by the condensation reaction between hydrazin sulfate and hexamethylenetetramine, forming an aqueous suspension of an insoluble white polymer (Ibrahim et al., 2017). This solution was purchased from Sigma-Aldrich (St. Louis, MO, USA) and it is a reference Material (CRM) that was produced and certified in accordance with ISO 17034 and ISO/IEC 17025 (Sigma-Aldrich, 2019).

DI water
DI water was filtered following the same methodology as mention in Section 2.1.1 for pH.

Brilliant Blue R dye solution
A dye solution of 5 mg/mL concentration was made by dissolving in DI water Brilliant Blue R dye with ca. 90% dye content and of molar mass 826 g/mol purchased from Sigma-Aldrich (St. Louis, MO, USA). This chemical has an absorption maximum at a wavelength between 560 and 575 nm (*Brilliant Blue R Product Information*).

Colloidal Solution
A colloidal solution with a concentration of 11 % v/v composed by Silica particles of an approximated size of 300 nm and ethanol of purity greater than 90 % from Sigma-Aldrich (St. Louis, MO, USA) was prepared by an external researcher.

2.2.2 Turbidity Prototype Sensor

The prototype is a multi-beam turbidity sensor, which is composed of two Light Emitting Diodes (LEDs) and two photo-transistors, and is most commonly used to measure low ranges (0-40 NTU) (Anderson, 2005). The range of interest was 0 to 10 NTU making this technology suitable for drinking-water applications.

Figure 1 shows a schematic of the sensor. It is composed of two concentric cylinders with a length of 80 mm and an inner and outer radii of 2 and 10 mm respectively. The LEDs and photo-transistors are arranged in a cross-shape configuration where each one is perpendicular from one another.

2.2.3 Wavelength Selection

Two LEDs, one of wavelength of 570 nm and the other of 850 nm, and two photo-transistors were used. The latter elements were purchased from Vishay Intertechnology Inc to have matching peak wavelengths with LEDs. The model BPW85 was used for photo-transistor 1, whose peak wavelength is 570 nm (Vishay Intertechnology Inc., 2019), whilst for photo-transistor 2 a TEPT4400 model with a peak wavelength of 850 nm (Vishay Intertechnology Inc., 2017).

850 nm

According to ISO7027 Standard (Wang et al., 2018), a wavelength between 830 and 890 nm should be selected to minimise the interferences light-absorbing substances could have on the accuracy of the sensor. Following this guideline, a wavelength of 850 nm was chosen, which is located in NIR. This LED was purchased from Würth Elektronik (Wurth Elektronik GmbH & Co. KG, 2019).

570 nm

As Hale et al. state, the absorption factor in water of 570 nm (0.00079 cm^{-1}) is two orders of magnitude lower than that of 850 nm (0.043000 cm^{-1}) (Hale et al., 1973). This suggests that less light will be absorbed by photo-transistor 1 compared to photo-transistor 2. Additionally, it is an uncommon wavelength to use for measuring turbidity, hence testing its effect on the prototype's performance is of interest. The 570 nm wavelength is green light located in the visible light region and it was obtained from Premier Farnell Limited (Premier Farnell Limited, 2019).

2.2.4 Measuring Principle

The sensor measures two different phenomena: nephelometry and turbidimetry.

Nephelometry follows the GLI Method 2, which is based upon a comparison of the intensity of light scattered by the sample under defined conditions with the intensity of light scattered by a standard reference suspension (Anderson, 2005). As the intensity of the scattered light increases, so does the turbidity. Particles interact with incident light by absorbing the light energy, which is then radiated in all directions. This re-radiated light is the scattering of incident light, which depends on particle size, shape and composition as well as the wavelength of the incident light.

The prototype measures side scattering which is the amount of light scattered at an angle of 90° from the light beam. This measuring angle minimises the effect of particle size on the measurement. Additionally, a simple optical design that is free from stray light can be used.

Turbidimetry determines the degree of light transmitted using a measuring angle of 180°. Transmission (T) is described by the Beer-Lambert law, as the ratio of transmitted light intensity (I) to light source intensity (I_o) which can be related to absorption using absorbance (A) as shown in Equation (1).

$$T = \frac{I}{I_o} = 10^{-A} \tag{1}$$

These two phenomena are detected by the photo-transistors. Photo-transistor 1 measures both the scattered light of the 570 nm LED and the transmitted light of the 850 nm LED. photo-transistor 2 measures the scattered light of the 850 nm LED and the transmitted light of the 570 nm LED.

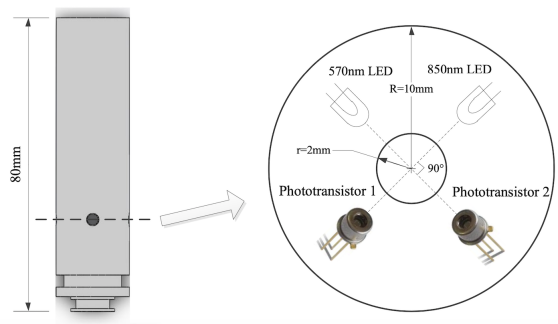


Figure 1: Schematic of the turbidity sensor prototype

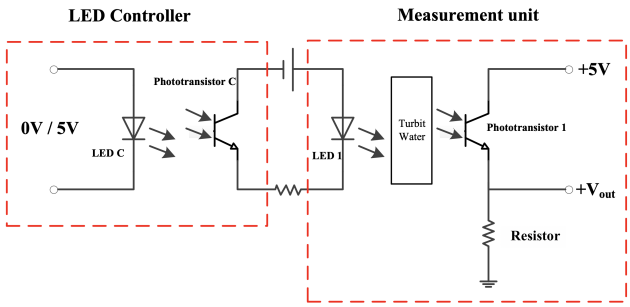


Figure 2: Schematic of the breadboard circuit

2.2.5 Sensor setup

The turbidity sensor prototype was connected to a breadboard circuit by eight wires. These wires were placed in pairs for each of the four elements (two LEDs and two light detectors). The breadboard circuit is mainly divided into two parts as shown in Figure 2: LED controllers and measurement unit. The LED controller is composed of a LED and a phototransistor, whose working principle is similar to a relay. Two Tenma 72-10500 power supplies provided constant power to the sensor's LEDs controlled via NI USB-6211, which is a multifunction data acquisition (DAQ) device that offers analog input/output, digital input, digital output, and two 32-bit counters (*USB-6211 Multifunction I/O Device*). NI USB-6211 gave a digital output signal of 0 V or 5 V. When the signal was 5 V, the current could go through both photo-transistor C and LED 2, thus the measuring unit started functioning.

In order to measure the desired quantities, phototransistors 1 and 2 needed to be supplied with a constant voltage of 5 V, hence an additional power supply IPS 303A from Iso-Tech was used. The light intensity shining on the light detectors was be calculated by collecting the voltage of the resistors. In the case of phototransistor 1, it was be done by measuring the voltage of the 10 kΩ resistor. For simplicity, only LED 1 and photo-transistor 1 were represented in Figure 2. LED 2 and photo-transistor 2, whose resistor has a constant value of 0.12 kΩ, followed the same circuit configuration.

The current and voltage of the desired properties constituted the breadboard circuit's output signals, which were then connected to a computer through NI USB-6211. By using LabVIEW, these output signals' data were acquired.

2.2.6 ABB Aztec ATS430 Sensor

This ABB sensor is constituted of two parts connected by EZLink: the sensor (ATS430) and the transmitter device (AWT420).

The sensor is suitable for turbidity and suspended solid measurements, however for the scope of this project, only the turbidity measurements were considered. The ATS430 sensor is able to function over a wide range of turbidity values (0-4000 NTU) and it has an auto-clean system to overcome optical fouling. The measuring principle used by this sensor is nephelometry. The LED emits light at a wavelength of 850 nm and side scattering is measured by the light detector (ABB, 2019). It complies with the ISO7027 Standard (Wang et al., 2018).

AWT420 is a four-wire single and dual-transmitter that is used for the measurements and control of parameters. In addition to turbidity measurements, this device can also be used to measure many other properties including pH and conductivity. It contains a removable SD card for data-logging.

3 Methodology

3.1 pH

3.1.1 Solution Preparation

Multiple buffer solutions with a pH value ranging from 5.9 to 8.1 with increments of 0.2 were prepared to test the sensors. Additionally, solutions at pH 6 and 7 were also prepared to validate their accuracy. These aqueous solutions were made by mixing KH_2PO_4 and NaOH as stated in (De Lloyd, 2000). All solutions were composed of 100 mL of DI water and KH_2PO_4 and NaOH of concentrations 0.1 M each. The required volume of NaOH varied depending on the desired pH value and it was calculated using De Lloyd's article as a reference (De Lloyd, 2000).

3.1.2 Sensor Calibration

All the sensors were calibrated using the three-point method. The calibrations were originally performed using the certified pH buffer solutions, whose pH values were 4, 7 and 10. The PyroScience pH miniprobe and Aquaread AP 2000 were calibrated using their corresponding software (Pyro pH Logger and AquaCal, respectively). In the case of the Microsens sensor, a calibration curve had to be constructed using the voltage data obtained from the LabVIEW program. Inaccurate behaviour was observed in the results of the PyroScience optical pH sensor when the certified pH buffers were used to calibrate it, thus a second calibration was performed using in-house buffer solutions at pH 5.9, 7 and 8.1.

Once each sensor was calibrated, measurements were taken to analyse their response time and accuracy.

3.1.3 Sampling Method

The appropriate software was loaded for each sensor and measurements were taken for every buffer solution prepared. The sensors were immersed in the so-

lutions and in the case of the optical ones, the beakers were covered in black cardboard to avoid any light disturbances. Once the sensor stabilised, the experiment ran for another 2 minutes before moving on to the next test. During the whole run, data was collected and the average stable pH value and the time taken to stabilise were then calculated.

3.2 Turbidity

3.2.1 Solution Preparation

Solutions ranging from 0 to 10 NTU in increments of 0.5 NTU were prepared by mixing DI water and the turbidity 1000 NTU calibration standard solution. A constant volume of 250 mL of DI water was used for every sample and 0.25 mL of the standard solution was added per NTU. Prior to producing the solutions, the calibration standard solution was shaken to produce a suspension. The in-house prepared NTU solution was then sonicated for 15 minutes to make the suspension more homogeneous, using a VWR Ultrasonic Cleaner USC-T. Subsequently, it was placed on a magnetic stirrer for the duration of the experiment. A 200 NTU solution was also made to calibrate the ABB meter following the same method.

The colloidal solution was then centrifuged in a Thermo Scientific Heraeus Megafuge 8 Centrifuge to separate particulates from the liquid phase. After decanting the liquid, the particles where weighted and mixed with DI water. Subsequently,sonication was carried out for the next 15 minutes to ensure a more homogeneous mixture was obtained. The approximate turbidity of the mixture was measured by the ABB Aztec ATS430 Sensor and a value of around 380 NTU was obtained. The solutions were diluted by adding more DI water to obtain a lower turbidity range, from 1 to 100 NTU. They were tested with the ABB meter to get an estimate of the turbidity, which was utilised when analysing them with the prototype. These solutions were also previously sonicated.

3.2.2 Sensor Testing

All of the prepared solutions were tested 10 times using the prototype sensor to evaluate how reproducible the measurements were. Before starting any measurements, the three power supplies were switched on and the electronics were connected.

A volume of 1 mL of solution was inserted in the sensor and two rubbers were used to cover the inlet and outlet. Additionally, tests were performed with the lights switched on and then off to determine the sensitivity of the prototype with stray light. The LabVIEW program was started and a 40 second measurement was taken per run. LED 1 was operational for the first 20 seconds, whilst LED 2 was switched on for the last 20 seconds.

Once the measurements were completed, the wires connecting the sensor to the breadboard were disconnected and the two rubbers covering the sensor were removed. The solution was then drained and a flow of dry air was used to eliminate the remaining bubbles or droplets inside the prototype sensor. The same procedure was followed for every sampling test.

In the case of the ABB Aztec ATS430 Sensor, a two-point calibration was first performed using a constant volume of 200 mL using filtered DI water (0 NTU) and a 200 NTU certified standard solution. Measurements took place inside the calibration pot (part no. ATS4000740), whose design minimised the effect of external light on the readings. After fully inserting the sensor into the pot, data logging was started until a stable reading was obtained. Subsequently, the meter was extracted, rinsed with water and dried thoroughly (ABB, 2019). The same steps were followed for every turbidity solution tested.

3.2.3 UV-Vis Spectroscopy

A Shimadzu UV-2600 spectrophotometer was used to obtain the absorption curves for a wavelength range of 200 to 900 nm. The 1000 NTU Calibration Standard solution and the in-house prepared Brilliant Blue R dye solution were tested. Turbidity solutions of 1, 2, 5, 10 and 20 NTU as well as a dye solution with a concentration of 0.5 mg/mL were prepared. Their stability was checked and the theoretical current of photo-transistor 1 with the dye solution was calculated.

By using the UVProve 2.70 Software developed by Shimadzu, a baseline correction was performed using filtered DI water as a reference. Measurements were taken by placing a cuvette filled with DI water in the rear cell holder and another one filled with the desired sample solution in the front cell holder with the optical faces facing the sides. Polymethyl methacrylate (PMMA) square cuvettes of 10 mm optical length from Brand Gmbh & CO. KG were used and hence the path length was assumed to equal 10 mm (Shimadzu Corporation, 2011). Three measurements were collected for each solution with a waiting interval time of approximately 5 minutes in between.

3.2.4 Current Determination

The theoretical current of photo-transistor 1 was calculated. The values obtained were compared to the current measured by the prototype due to transmission.

Theoretical Current Determination

BPW85 photo-transistor's data sheet (Vishay Intertechnology Inc., 2019) and the 570 nm current and transmission measurements collected by LabVIEW were used. There was no data available for the BPW85 model, hence the BPW85A model was used to perform the calculations as they follow similar trends. Using Luminous Intensity vs. Forward Current on the BPW85 data sheet, the current (C) was converted to luminous intensity (I_v) via Equation (2), which was then converted to $\frac{Lumen}{Area}$ using equations (3) obtained from (*LED mcd (millicandela) to Lumens Converter*) and (4), respectively.

$$I_v = 0.1067 \times C - 0.2833 \tag{2}$$

$$Lumen\ (lm) = I_v\ (mcd) \times \frac{[2\pi(1 - cos(\frac{\pi\theta}{360})]}{1000} \tag{3}$$

$$\frac{Lumen}{Area}\ (\frac{lm}{cm^2}) = \frac{Lumen}{\pi \times r^2} \tag{4}$$

where θ is twice the viewing angle, thus a constant value of 32° was used (Premier Farnell Limited, 2019)

and the radius (r) of the sensor was set to 2 mm as represented in Figure 1.

Using Equation (5) obtained from (*Lumen/square centimenter to Watt/square centimeter Converter*), irradiance (E_e) was determined, which corresponds to the light source intensity (I_o) in the Beer-Lambert Equation (1).

$$E_e\,(\frac{W}{cm^2}) = \frac{Lumen}{Area} \times 0.00146 \tag{5}$$

The transmitted light intensity (I) can be calculated from Equation (1), where the absorbance (A) obtained from the Shimadzu UV-2600 experiments. To compare the theoretical value to the experimental one, the transmitted light intensity was converted to current (C_{th}) using Equation (6) and Figure 4 from the BPW85 data sheet (Vishay Intertechnology Inc., 2019). The graphs were plotted for a reference wavelength of 950 nm , to account for the relative spectral sensitivity. Equation (6) had to be corrected using a correction factor of ($\frac{0.3}{0.7}$), as stated in Figure 10 from the BPW85 data sheet (Vishay Intertechnology Inc., 2019).

$$C' = 1.1549 \times E_e - 0.0069 \tag{6}$$

$$C_{th} = \frac{0.3}{0.7} \times C' \tag{7}$$

Experimental Current Determination

Ohm's law (8) was used to calculate the current (C_m) measured by the resistance due to transmission of light-detector 1. The resistor had a constant value of 10 kΩ and the voltage values were measured by LabVIEW.

$$C_m\,(mA) = \frac{Voltage}{Resistance} \tag{8}$$

4 Results & Discussion

4.1 pH

The three different commercial sensors were compared against each other to determine which one was the most accurate and the fastest to stabilise.

4.1.1 Accuracy

The relative error was calculated for every solution and sensor tested using the Mettler Toledo pH meter as the reference. The errors were then plotted against the pH values measured by the reference sensor as shown in Figure 3. The sensor manufactured by PyroScience showed very accurate behaviour, where the relative error was between 0 and 2 %. The Aquaread's sensor followed a similar trend, with its relative error slightly higher ranging from -1 to 2 %.

The Microsens sensor had significantly higher relative errors (between 10 and 18 %) than the previous two. This could have been due to the fact that the sensor was calibrated using certified pH buffer solutions which tend to not always be chemically stable. Hence, a second run was performed, where the in-house buffer solutions at pH 5.9, 7 and 8.1 were used to calibrate the sensor. The same methodology as the previous runs was followed and the results are plotted on Figure 3. By using the in-house buffer solutions, the relative error of Microsens sensor was reduced to 10 % or lower values. However, the error was still considerably larger compared to the others. The integrated electronics of the sensor could be the potential cause of its inaccurate behaviour. Overall, the PyroScience optical pH sensor was the most precise and accurate.

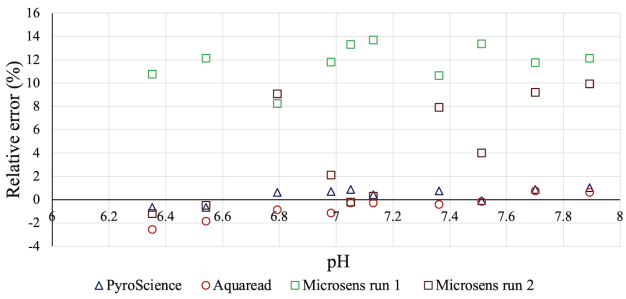


Figure 3: Relative errors of tested commercial pH sensors with respect to the reference meter

4.1.2 Response Time

A sensor not only has to be as accurate as possible but also has to be fast to provide a stable measurement, thus the time taken to stabilise was measured. These values are summarised in Table 2, where it can clearly be seen that PyroScience Optical sensor was the quickest to reach a constant value, followed by Aquaread. Microsens took 8 more minutes compared to the optical probe.

Table 2: Stabilising time of evaluated pH sensors

Sensor Model	Time (min)
PyroScience	2
Aquaprobe	5
Microsens	10

4.1.3 Testing Different PyroScience pH Miniprobe Units

From Sections 4.1.1 and 4.1.2, it could be concluded that PyroScience Optical pH sensor was not only the most accurate but also the fastest to stabilise. However, only one unit of the product had been tested. Thus, the same experiments were conducted for a second device to analyse if these properties were consistent throughout different units. In Figure 4, the relative errors for both were plotted. The two sensors gave very similar trends, with most of their relative being errors below 1 %. Additionally, the time response of the second unit was 2.5 minutes, only half a minute more than for the first one. Hence, it can be stated that the different probes shared similar accuracy and both reached steady state quicker than their competitors.

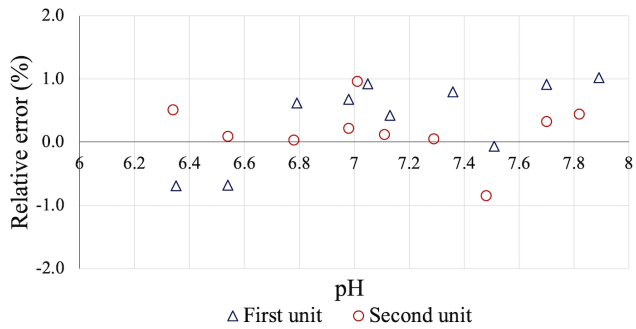


Figure 4: Relative errors of different PyroScience units with respect to the reference meter

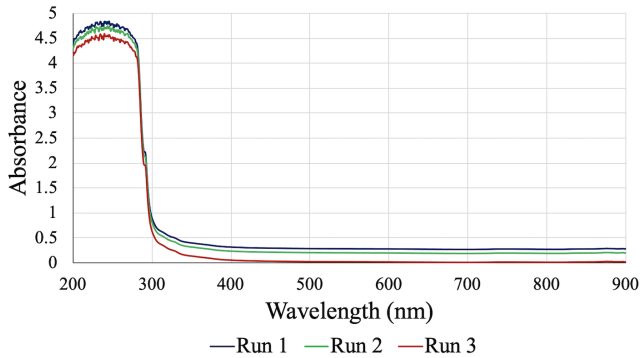


Figure 5: Absorption spectra from 200 to 900 nm wavelength for a 10 NTU turbidity solution with 5 minutes difference between runs

4.2 Turbidity

4.2.1 Transmission and Scattering using Standard Solutions

The voltage measured due to transmission is shown on Figures 6a and 6c, whilst the voltage measured due to scattering is plotted on Figures 6b and 6d for solutions with calculated turbidity values of 1 and 10 NTU.

In the case of transmission, voltage should decrease as turbidity increases. As the amount of suspended materials in the solution increase, the transmission of light is progressively decreased until it is inhibited. However, this trend cannot be observed in Figure 6. Moreover, the 850 nm LED gave a constant voltage of 5 V regardless of the run number or solution's turbidity. The 570 nm exhibited a erratic trend throughout the runs, which is unexpected as it should measure similar values to be considered a stable measurement. Like the 850 nm LED, the 570 nm one showed no significant changes in voltage as turbidity changes. Values in a 2 to 4 V range were measured for all solutions.

For scattering, the voltage should increase as turbidity increases. A similar issue was encountered when performing the measurements, where the 570 nm LED gave a constant voltage of approximately 0.25 V and the 850 nm one had an erratic behaviour for all solutions tested.

Transmission of the 850 nm and scattering of the 570 nm LEDs exhibited the same trend, where as turbidity changes, the voltages stayed constant. These two properties were measured by photo-transistor 1. Conversely, photo-transistor 2 measured the transmission of the 570 nm and 850 nm LEDs, which displayed an unsteady behaviour where no clear changes in voltage could be observed. This suggests that the photo-transistors might not be sensitive enough to detect the changes in turbidity. Furthermore, it could be affirmed that each photo-transistor displays a similar trend throughout the experiments. Another possible reason for the inaccuracy of the data could be that the Sigma-Aldrich standard solutions used were not chemically stable, hence further testing was done.

4.2.2 Stability check of the Standard Solution

To check whether the turbidity standard solution was stable, the Shimadzu UV-2600 spectrophotometer was used due to having a similar path length as the prototype. The results obtained for a turbidity value of 10 NTU are shown in Figure 5. For every solution the absorbance decreased, as time increased, implying that the solutions were unstable.

Therefore, a chemically stable compound was needed to effectively study the performance of the prototype. The same experiments were conducted using Brilliant Blue R dye and the prepared colloidal solution. It was determined that they were chemically stable, thus the subsequent experiments were performed using these solutions.

4.2.3 Performance check of Photo-transistor 1

The theoretical and experimental currents of photo-transistor 1 were calculated using the method stated in Section 3.2.4 for the dye solution of concentration 0.5 mg/mL. The results are summarised in Table 3.

Photoransitor 1 is a BPW85 model, but no data was available for it. Therefore, the data of a similar model (BPW85A) was used to perform the theoretical calculations, which led to deviations in the results.

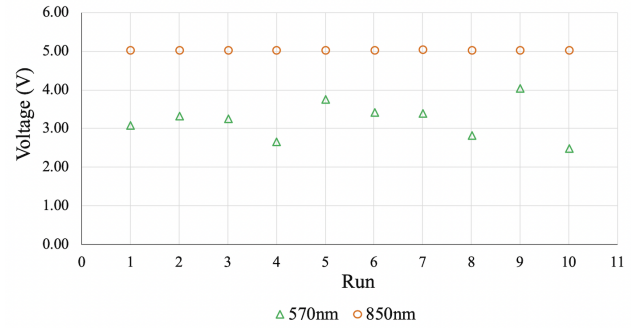
As the solution was stable, the absorbance at a wavelength of 570 nm was similar throughout the runs, hence the theoretical current (C_{th}) gave a constant value of approximately 0.0029 mA. Although both parameters are within the same order of magnitude, the measured current (C_m) gave an absolute maximum relative error of 31%. Therefore, the data presented suggests that photo-transistor 1 is not only inaccurate but also inconsistent. In an attempt to overcome this obstacle, the sensor and testing conditions were modified as explained in Section 4.2.4.

Table 3: Theoretical and experimental current summary for Brilliant Blue R dye solution

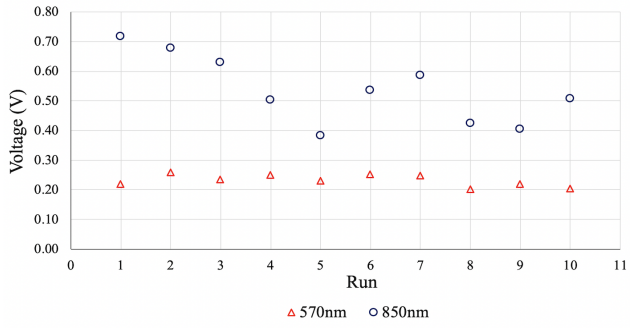
Run	C_{th} (mA)	C_m (mA)
1	0.0029	0.0025
2	0.0028	0.0019
3	0.0029	0.0021

4.2.4 Study of light effects on measurements

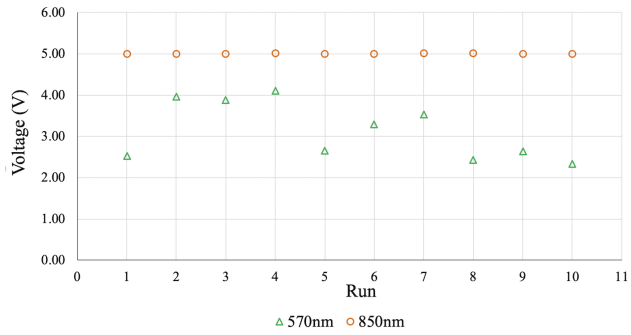
The effect that ambient light could have on the accuracy of the measurements was tested by performing experiments with the lights on and then off. The relative errors with respect to the results obtained from the ones with the lights on were calculated. These values are summarised in Table 4. Both phototransistors were severely affected by this light.



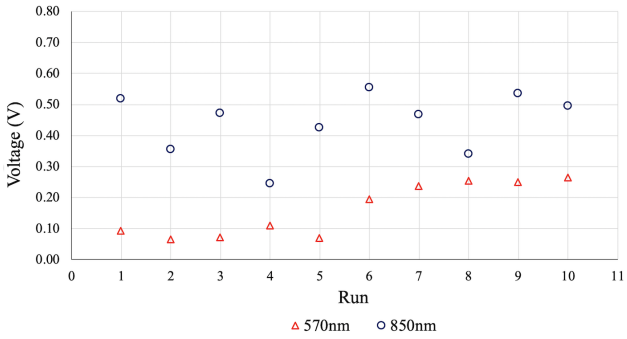
(a) Transmission 2 NTU



(b) Scattering 2 NTU



(c) Transmission 10 NTU



(d) Scattering 10 NTU

Figure 6: Transmission and scattering plots of both LEDs using 2 NTU and 10 NTU turbidity solutions for 10 experimental runs

The relative errors due to the transmission of light from both LEDs were between the same range. In the case of scattering due to the 850 nm LED, they varied from -31 to 15 % and for scattering of the 570 nm, maximum and minimum relative errors of -63 and -3 % were observed. Despite the significant differences in the voltage measured due to this phenomenon, photo-transistor 1 was more affected than photo-transistor 2. The former being more susceptible to larger fluctuations when the lights are on as its peak wavelength is located in the visible light range. Therefore, the ambient light should be minimised when performing experiments.

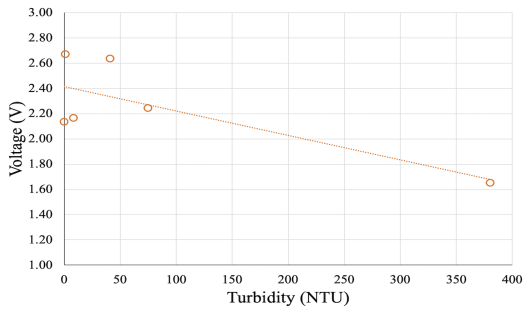
Table 4: Relative errors measured for transmission and scattering with the lights on and off for photo-transistor 1

	Transmission (V)	
NTU	570 nm	850 nm
0.9	16.479	-6.834
8.5	-2.117	17.171
41	-14.022	-14.662
75	-10.279	-8.874
380	-7.271	-6.614
	Scattering (V)	
NTU	570 nm	850 nm
0.9	-32.089	6.114
8.5	-61.524	-30.805
41	-54.442	-19.689
75	-3.221	-21.978
380	-18.689	-14.592

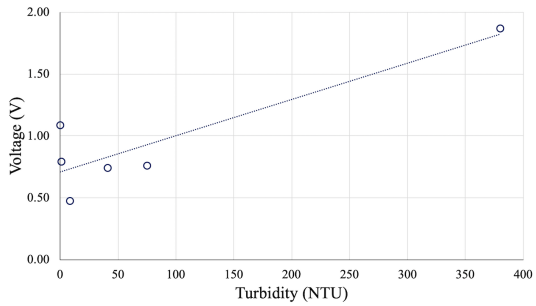
4.2.5 Prototype performance using the colloidal solution

The last set of experiments was conducted using the colloidal solution following the recommendations stated in Section 4.2.4.

In Figure 7, the expected trends for transmission and scattering due to the 850 nm LED were noticeable for the larger turbidity range. However, the lower range still showed an irregular and non-linear tendency. Similar trends were observed for the 570 nm LED. Although using a chemically stable solution did improve the behaviour of the sensor, it was still not able to accurately measure the turbidity in the range of interest for drinking-water applications, which is between 0 to 10 NTU.



(a) Transmission of the 850 nm LED



(b) Scattering of the 850 nm LED

Figure 7: Transmission and scattering plots of the 850 nm LED using 5 dilutions of the colloidal solution with approximate turbidity values of 0.9, 8.5, 41, 75 and 380 NTU. Dashed lines are linear fits to the data

5 Conclusion

The study has successfully determined the most suitable pH technology to implement in the MFMS prototype. PyroScience pH miniprobe was not only the most accurate but also the fastest to reach a stable measurement. These properties in combination with its consistent behaviour throughout the different tested units and its reduced size make it the most suitable pH sensor. There is currently a commercial version of this product nevertheless, it must be taken into consideration that the ones evaluated were prototypes.

The implementation of a turbidity sensor was more challenging. It was shown that the certified standard turbidity solutions were unstable and that in-house ones should be prepared. Additionally, the ambient light had a significant impact on the measurements. Thus its minimisation is recommended, especially in the case of scattering due to the 570 nm LED. As turbidity increased, downward and upward trends were observed for transmission and scattering, respectively. This was noticeable for high values of turbidity yet the photo-transistors were unable to detect changes in the lower range. It can therefore be concluded that it is necessary to optimise the design of the turbidity prototype sensor and that high accuracy elements need to be used. An optical lens should be placed after the LEDs and photo-transistors to minimise the viewing angle and ease the cleaning process. The diameter of the inner cylinder should be larger to increase the path length. The use of specific electronic elements will reduce the noise and increase the signal-to-noise ratio (SNR). By amplifying the output voltage and current signals given by the photo-transistors, clearer changes may be observed. The use of different LEDs as well as light detectors, such as the BPW 85C model, may lead to more accurate results. Once the design is optimised, the prototype's performance should be re-evaluated.

6 Acknowledgements

We would like to thank Dr Chao Yuan, Parth and everyone in the SPEL Group for their invaluable support and guidance throughout this project.

7 References

ABB (2019). *ABB Aztec ATS430 Turbidity sensor*. Tech. rep., p. 27. URL: https://library.e.abb.com/public/bd3d87c0cd73498c8a3c4166de4eeb3e/OI_ATS430-EN_H.pdf.

Anderson, C.W. (2005). "Techniques of Water-Resources Investigations". In: *9*. Chap. A6.

Aquaread. *AquaCal – Aquaprobe Live Data & Calibration Utility with Report Generation*. URL: https://www.aquaread.com/portofolio/aquacal-2/ (visited on 12/10/2019).

De Lloyd, D. (2000). *Preparation of pH buffer solutions*. URL: http://delloyd.50megs.com/moreinfo/buffers2.html (visited on 12/07/2019).

Endmemo. *Lumen/square centimenter to Watt/square centimeter Converter*. URL: http://www.endmemo.com/sconvert/lumen_squarecentimeterwatt_squarecentimeter.php (visited on 12/11/2019).

Hale, George M. et al. (Mar. 1973). "Optical Constants of Water in the 200-nm to 200-μm Wavelength Region". In: *Applied Optics* 12.3, pp. 555–563. DOI: 10.1364/AO.12.000555. URL: https://www.osapublishing.org/abstract.cfm?URI=ao-12-3-555.

Ibrahim, S. F. et al. (May 2017). "Metrological Evaluation of Prepared Turbidity Working Standard Solutions to achieve Traceability for Textile Waste Water Effluents". In: *Egyptian Journal of Chemistry* 0.0, p. 13. DOI: 10.21608/ejchem.2017.938.1046. URL: https://ejchem.journals.ekb.eg/article_3291.html.

Idronaut S.r.l (2019). *Optical Oxygen Sensor - Idronaut*. URL: https://www.idronaut.it/sensors/optical-oxygen-sensor/ (visited on 12/10/2019).

Johnson, Nora H. (Sept. 1961). "Drinking-Water Standards". In: *Nature* 191.4795, pp. 1261–1261. ISSN: 0028-0836. DOI: 10.1038/1911261a0. URL: http://www.nature.com/articles/1911261a0.

Kitchener, B. GB et al. (2017). "A review of the principles of turbidity measurement". In: *Progress in Physical Geography: Earth and Environment* 41.5, pp. 620–642. DOI: 10.1177/0309133317726540. URL: http://journals.sagepub.com/doi/10.1177/0309133317726540.

Kremesti, R. E. *Glass pH electrode vs ISFET pH Electrode*. URL: http://kremesti.com/water/glass_vs_ISFET_pH_Electrode.htm (visited on 12/07/2019).

Mettler Toledo (2013). "pH Theory Guide". In: p. 99. URL: https://www.mt.com/dam/non-indexed/po/pro/Brochures/IB_pH_Theory_Guide_EN_230113.pdf.

Microsens SA (2017). *MSFET 3332*. Tech. rep., p. 7. URL: www.microsens.ch.

National Instruments. *LabVIEW 2019*. URL: https://www.ni.com/es-es/shop/labview/labview-details.html.

– *USB-6211 Multifunction I/O Device*. URL: https://www.ni.com/es-es/support/model.usb-6211.html.

Ohms Law Calculator. *LED mcd (millicandela) to Lumens Converter*. URL: http://www.ohmslawcalculator.com/mcd-to-lumens (visited on 12/11/2019).

Paragon Scientific Ltd (2017). *Certified pH Buffer Solutions*. Tech. rep., p. 2. URL: https://www.paragon-sci.com/pub/media/productattach/c/e/certified_ph_buffer_solutions_flyer.pdf.

Peacock, C. (2014). *Multiparameter AP-2000 - Water Analysis Aquaprobes*. URL: https://www.aquaread.com/portofolio/ap-2000/ (visited on 12/07/2019).

Premier Farnell Limited (2019). *MCL053SGC - LED, Green, Through Hole, T-1 3/4 (5mm), 20 mA, 2.1 V, 572 nm*. URL: https://il.farnell.com/multicomp/mcl053sgc/led-5mm-16-super-green/dp/1581182 (visited on 12/08/2019).

PyroScience GmbH. *pH Sensor Manual - Piccolo2-OEM*.

– *Piccolo2 Fiber-Optic Oxygen Meter User Manual*.

– *Piccolo2-OEM Fiber-Optic Oxygen Meter OEM-Module with UART-Interface Manual*.

– (2019). *Robust Optical pH Probes*. URL: https://www.pyroscience.com/robust-optical-pH-probes.html (visited on 12/07/2019).

Shimadzu Corporation (2011). *Shimadzu UV-Visible Spectrophotometer UV-2600/2700 Instruction Manual*. Tech. rep., p. 144.

Sigma-Aldrich. *Brilliant Blue R Product Information*. Tech. rep., p. 2.

Sigma-Aldrich (2019). *Turbidity 1000 NTU Calibration Standard - Formazin certified reference material — Sigma-Aldrich*. URL: https://www.sigmaaldrich.com/catalog/product/sial/turb1000?lang=es{\&}region=ES (visited on 12/08/2019).

Thermo Fisher Scientific. *Buffer Solution, pH 10.00, Color-Coded Blue (Certified)*. URL: https://www.fishersci.com/shop/products/buffer-solution-ph-10-00-color-coded-blue-certified-fisher-chemical-4/sb11520{\#}?keyword=Buffer+Solution+pH+10.00+Color-Coded+Blue+Certified+Fisher+Chemical (visited on 12/10/2019).

– *Buffer Solution, pH 4.00, Color-Coded Red (Certified)*. URL: https://www.fishersci.com/shop/products/buffer-solution-ph-4-00-color-coded-red-certified-fisher-chemical-4/sb1014 (visited on 12/10/2019).

– *Buffer Solution, pH 7.00, Color-Coded Yellow (Certified)*. URL: https://www.fishersci.com/shop/products/buffer-solution-ph-7-00-color-coded-yellow-certified-fisher-chemical-5/sb1074 (visited on 12/10/2019).

United Utilities (2019). *pH*. URL: https://www.unitedutilities.com/help-and-support/your-water-supply/about-your-water/ph (visited on 12/07/2019).

U.S. Geological Survey. *pH and Water*. URL: https://www.usgs.gov/special-topic/water-science-school/science/ph-and-water?qt-science{_}center{_}objects=0{\#}qt-science{_}center{_}objects (visited on 12/10/2019).

Vishay Intertechnology Inc. (2017). *TEPT4400*. URL: https://www.vishay.com/docs/81341/tept4400.pdf.

– (2019). *BPW85, BPW85A, BPW85B, BPW85C*. URL: https://www.vishay.com/docs/81531/bpw85a.pdf (visited on 12/08/2019).

Wang, Y. et al. (June 2018). "Low-Cost Turbidity Sensor for Low-Power Wireless Monitoring of Fresh-Water Courses". In: *IEEE Sensors Journal* 18.11, pp. 4689–4696. DOI: 10.1109/JSEN.2018.2826778. URL: https://ieeexplore.ieee.org/document/8337739/.

Williams, D. R. (Imperial College London) (2018). "Multi-Functional Micro Sensors (MFMS) for Process Control". In:

Wurth Elektronik GmbH & Co. KG (2019). *15400385A3590 Datasheet(PDF)*. URL: https://www.alldatasheet.com/datasheet-pdf/pdf/940498/WURTH/15400385A3590.html (visited on 12/08/2019).

Taylor bubble characteristics in curved microfluidic channels

Hon Wa Yew, Mingrou Xie

Imperial College London, Department of Chemical Engineering, Imperial College Rd, Kensington, London SW7 2AZ

12th December, 2019

Abstract. Understanding of bubble characteristics is a key step in developing predictive models for two-phase flows in microfluidics. In this work, several bubble characteristics are investigated for a single Taylor bubble in a curved microfluidic channel at Reynolds number much lower than 10, where surface tension effects dominate, and at capillary numbers of 0.082 and 0.0082. An expression relating two-phase velocity and bubble velocity was derived for ease of transitioning from experiments to analysis of bubble characteristics and finds good agreement at capillary numbers of 0.082. The asymmetry in film thickness and bubble head position characteristic of bubbles in curved channels is also profiled using numerical results from 2-D computational fluid dynamic models using Volume of Fluid (VOF) interFoam solver from an open-source finite-volume software OpenFOAM®. With increasing channel curvature, bubble length increases by as much as 50%, with implications for increased surface area-to-volume ratios. Finally, the pressure drop across the bubble was found to not be affected by the presence of a curved channel. The methodology developed here can be used in future work to investigate Taylor bubble flows in 3-D curved channels and other interesting geometries.

1 Introduction

Work on bubbles in microfluidic channels has seen increased attention in recent years due to rapidly expanding interest in microfluidics for numerous applications [1]. Numerous bioengineering technologies utilise microfluidics as it makes use of small fluid volumes and length scales typical in biology [2, 3]. Novel microfluidic techniques involving bubbles have been explored, which employ bubbles as micro actuators [4], micro mixers [5], or compartmentalised micro reactors [6]. On the other hand, undesired bubbles that form and adhere to walls due to high surface tension can increase flow resistance and hence pressure drop across the channel, which can be unfavourable where precise control of fluid flow is critical, as is often the case in microfluidic applications [7]. In cell culture and transport, unwanted bubbles also increase shear stress on the cells, causing damage and cell death [8].

In both scenarios, a thorough comprehension of the geometry and other characteristics of bubbles in micro-scale channels would provide better understanding about how they affect flow and pressure drops. To this end, typical characteristics of interest that have been studied include film thickness, bubble head and tail curvature, bubble velocities and pressure drop across bubbles [9–11], for Taylor bubbles which are of interest here.

Fairbrother and Stubbs first determined experimentally that a Taylor bubble travels slightly faster than the average fluid speed [12], and that a stagnant film of liquid surrounds the moving bubble [13]. The film thickness is of significance where mass and heat transfer to and from the channel walls are important. The presence of a liquid film due to Taylor bubbles also has other implications such as incomplete drainage when gas is used to push out the resident fluid.

In order to model how the film thickness changes with operating conditions and channel geometries, relationships between the former and latter would be useful. Bretherton first developed an analytical expression to characterise the film thickness in straight channels as a function of capillary number, initially for capillary numbers less than $5 \cdot 10^{-3}$ [14]. His equation was then refined empirically by Aussillous and Quere [11], whose expression was in turn given a theoretical justification by Klaseboer [10]. Cherukumudi made use of this methodology to derive a more sophisticated model that involves modelling both the front and the back of the bubble [9]. Separately, recent attempts have been made to characterise the bubble shape in both 2-D and 3-D straight channels [15, 16]. Characterisation of Taylor bubbles in straight channels appears to be well developed.

The field for curved microfluidic channels presents a different story. Some work has been done fairly recently on extending Bretherton's and Aussillous film thickness expression to bubble in

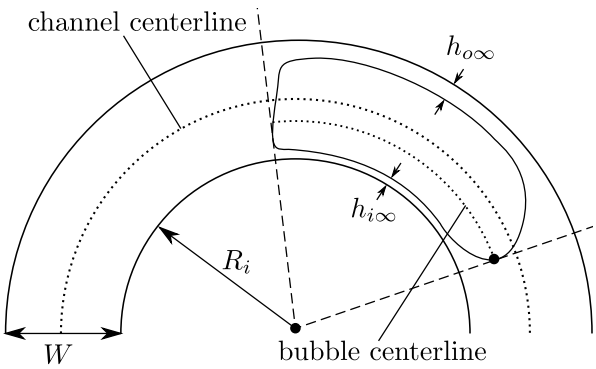


Figure 1: Schematic of CFD model. For bubble length, the length of the channel center line that lies between the dashed lines marking the sector where the bubble lies is taken as the bubble length.

curved channels. Muradoglu (2007) derived analytically an expression which showed that the inner film thickness of a Taylor bubble in a curved channel is thinner than the outer film thickness and that this asymmetry is the cause of higher bubble velocities around the bend [17]. His work sees immediate relevance as microfluidic channels are often curved in order to control fluid flowrates and to save on microfluidic chip areas (typical channel patterns include serpentine, zig-zag or square wave channels [18]).

However, a complete characterisation of Taylor bubble geometry in curved channels has yet to be done. In this paper, numerical simulations were carried out using an open-source computational fluid dynamics (CFD) software OpenFOAM®. The bubble velocity, film thicknesses, bubble length, bubble head position and pressure drop across bubble were investigated. All simulations were done at a fixed two-phase velocity (more experimentally accessible), so a relationship between bubble velocity and two-phase velocity was also derived.

2 Methodology

Numerical simulations were performed in 2-D using a Volume of Fluid with Flux-Corrected Transport (VoF-FCT) solver, interFoam, from OpenFoam® Version 7.0. The solver solves for two incompressible, isothermal and immiscible fluids using a one-fluid formulation through the introduction of a phase indicator function, α, where the volume fraction of a phase can be found by integrating it over a cell $\int_V \alpha \, dV$. α is equal to 1 for water.

For the incompressible system, the continuity equation is expressed as

$$\nabla \cdot \mathbf{u} = 0, \tag{1}$$

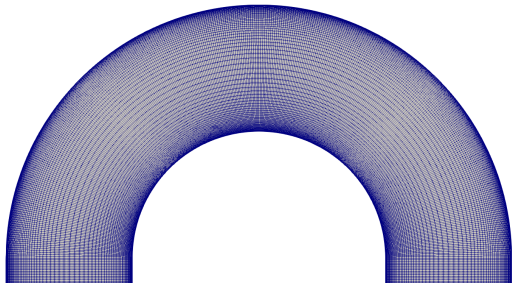


Figure 2: Mesh density of CFD model. There is a higher density of computational cells in the channel walls to allow for better resolution of film thickness.

and the momentum balance is

$$\frac{\partial(\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u}\mathbf{u}) = -\nabla p + \nabla \cdot \mu(\nabla \mathbf{u} + \nabla \mathbf{u}^T) + \mathbf{F}_\sigma, \tag{2}$$

where the force due to the surface tension, $\mathbf{F}_\sigma$ is obtained from the CSF Model [19].

$$\mathbf{F}_\sigma = -\sigma \nabla \cdot \left(\frac{\nabla \alpha}{|\nabla \alpha|} \right) \nabla \alpha \tag{3}$$

Transport of the interface is modelled using interface advection with bounded compression through the introduction of an artificial compression term, which is solved algebraically using the Multi-dimensionsal Limiter for Explicit Solution (MULES) algorithm [20].

$$\frac{\partial \alpha}{\partial t} + \nabla \cdot (\alpha \mathbf{u}) + \nabla \cdot (\alpha(1-\alpha)\mathbf{U}_r) = 0, \tag{4}$$

with a relative velocity term, $\mathbf{U}_r$, defined by

$$\mathbf{U}_r = C_\gamma \left| \frac{\phi}{|S_f|} \right| \mathbf{n}_f, \tag{5}$$

where ϕ is the volumetric flux, S_f is the surface area of the cell, n_f is the unit vector normal to the cell surface, and C_γ being a tuning parameter to determine the amount of compression. The $\alpha(1-\alpha)$ ensures that the term is only active at the two-phase boundary. The value of C_γ can range from 0 (no compression) to 4 (maximum compression). Higher compressions make the interface more sharply defined, but induces spurious currents as a tradeoff [21], leading to unphysical results. In our simulations, the value of C_γ is 1.

The density and viscosity are expressed as a weighted sum of the densities and viscocities of each fluid.

$$\rho = \alpha \rho_1 + (1-\alpha)\rho_2 \tag{6}$$

$$\mu = \alpha \mu_1 + (1-\alpha)\mu_2 \tag{7}$$

Figure 2 depicts a representative computational domain consisting of a channel of width w, with

a portion of straight channel before and after the curved channel with an inner radius of curvature, R_i. The dimensionless curvature $\beta = w/R_i$ is set to values of 0.1, 0.25, 0.33, 0.5, 1 and 1.33. This range is capped by a upper bound where a sufficiently long bubble can no longer fit within the length of the curved channel. A bubble is patched near the inlet of a long channel and is allowed to flow through sufficient distance before entering the domain of interest in order to ensure that the flow is fully developed. It was found that the the there is a minimum bubble length of about three times the width of the channel in order for it to have a noticeable constant film thickness in a region away from the head an tail of the bubble. In our setup, the bubble was patched to be about $3.5w$ in length, and the volume was kept constant for all runs.

Under consideration of computational expenses, mesh grading was used to reduce the number of cells in the channel center, yet maintain adequate mesh fineness near the channels in order to better resolve film thickness. After considerable mesh testing a geometric ratio of 0.92 between the center and wall cell was used. Here we define the channel center line as the line that runs through the middle of the whole channel. The domain of interest has 72 cells along the channel width. In the axial direction, the cell interval was kept at a constant value of $0.0125w$ along the centreline of the curved channel.

In this study, flow at two-phase capillary numbers $Ca_{TP} = 0.082$ and 0.0082 and a low Reynolds number Re of 0.1 were considered for model validation with Muradoglu's 2007 paper [17]. At low $Re < 10$, inertial effects can be ignored relative to viscous forces [11]. The density ratio and the kinematic viscosity between the two fluids were kept at 10^2. A fixed two-phase velocity $U_{TP,\text{inlet}}$ was set at the inlet for all cases and is set to zero gradient at the outlet. The pressure of the outlet is fixed at 0 at the outlet and is zero gradient at the inlet. The surface tension between the two fluids, σ is allowed to vary in order to achieve the desired two-phase capillary number.

The bubble geometry is quantified by taking sampling lines orthogonal to the channel walls. This effectively parameterises the domain into a set of curvilinear coordinates that specify position by the distance along the channel center line and the distance from the outer channel wall. Each sampling line contains 500 linearly-spaced points, and the interval between each sampling line is $0.0125w$. This allows for easy and automated post-processing for the determination of geometric properties of the bubble such as the film thicknesses and bubble length.

All of the bubble characteristics are governed by the bubble capillary number, $Ca_B = \mu U_B/\sigma$, rather than the two-phase capillary number, $Ca_{TP} = \mu U_{TP}/\sigma$, so a relationship between the two velocities is necessary in order to make meaningful comparisons with the literature. Abadie has formulated an expression for U_B as a function of Ca_{TP} for 2-D axisymmetric flow [22], which we take as a framework to adapt with Muradoglu's expression for film thicknesses in curved channels, to obtain a corresponding relationship with different values of β.

Expressions for the inner and outer film thicknesses, $h_{i\infty}$ and $h_{o\infty}$ have been derived by Muradoglu [17] using lubrication theory.

$$\frac{h_{i\infty}}{w} = \frac{1}{2+\beta}\left(\frac{2}{2+\beta}\right)^{2/3}\left(1.3375 Ca_B^{2/3}\right) \tag{8}$$

$$\frac{h_{o\infty}}{w} = \frac{1+\beta}{2+\beta}\left(\frac{2+2\beta}{2+\beta}\right)^{2/3}\left(1.3375 Ca_B^{2/3}\right), \tag{9}$$

Where Ca_B is the bubble capillary number,

$$Ca_B = \frac{\mu U_B}{\sigma} \tag{10}$$

Since there is a constant film thickness at a region of the bubble, we can obtain an expression for the slip velocity, W, by performing a mass balance, which is found to be in the familiar Bretherton form.

$$W = 1 - \frac{U_{TP}}{U_B} = \kappa Ca_B^{2/3}, \tag{11}$$

where

$$\kappa = 1.3375\left[\frac{2^{5/3}(1+(1+\beta)^{8/3}}{(2+\beta)^{8/3}}\right] \tag{12}$$

It is useful to express Ca_B as a function of Ca_{TP}. Here, we follow Abadie's approach [22] to derive a first order relationship between U_B and U_{TP}. Equation 11 can be written as

$$\frac{U_B}{U_{TP}} = \frac{1}{1-\kappa Ca_B^{2/3}} \tag{13}$$

At very low capillary numbers, U_B approaches U_{TP}, and Ca_B approaches Ca_{TP}, therefore

$$\frac{U_B}{U_{TP}} \sim 1 + \kappa Ca_{TP}^{2/3} + \kappa^2 Ca_{TP}^{4/3} + ... \tag{14}$$

Disregarding the higher order terms, we arrive at the expression

$$U_B = U_{TP}\left(1 + \kappa Ca_{TP}^{2/3}\right) \tag{15}$$

Equation 11 can now be written in terms of the two-phase capillary number.

$$W = \kappa Ca_{TP}^{2/3}\left(1 + \kappa Ca_{TP}^{2/3}\right)^{2/3} \tag{16}$$

Since the slip velocity has an upper bound of about 0.4 due to increasing inertial effects that cause it to deviate from the Bretherton behaviour. We may apply the Aussillous and Quere correction to find a more accurate relationship between expression for the slip velocity [11].

$$W = \frac{\kappa Ca_{TP}^{2/3}\left(1+\kappa Ca_{TP}^{2/3}\right)^{2/3}}{1+2.5\kappa Ca_{TP}^{2/3}\left(1+\kappa Ca_{TP}^{2/3}\right)^{2/3}}, \tag{17}$$

A fitting parameter, γ may be introduced in the Bretherton form (Equation 11) so that the final expression becomes

$$W = \frac{\kappa Ca_{TP}^{2/3}\left(\gamma+\kappa Ca_{TP}^{2/3}\right)^{2/3}}{1+2.5\kappa Ca_{TP}^{2/3}\left(\gamma+\kappa Ca_{TP}^{2/3}\right)^{2/3}} \tag{18}$$

This equation may be rearranged into a form where the dependence of U_B on U_{TP} is made explicit.

$$\frac{U_B}{U_{TP}} = \left[1-\frac{\kappa Ca_{TP}^{2/3}\left(\gamma+\kappa Ca_{TP}^{2/3}\right)^{2/3}}{1+2.5\kappa Ca_{TP}^{2/3}\left(\gamma+\kappa Ca_{TP}^{2/3}\right)^{2/3}}\right]^{-1} \tag{19}$$

In order to calculate the velocity of the bubble, two different methodologies were employed. For a bubble moving in a straight channel, the velocity of the centre of mass is a good way to quantify the velocity of the bubble. However, in a curved channel, the centre of mass travels a shorter distance compared to the bulk of the bubble, such that the bubble appears to have a lower magnitude of velocity than at the straight channel. However, assuming that the centre of mass of the bubble travels at a constant angular velocity, ω, we are able to find the tangential velocity of the bubble at the center line of the channel by

$$U_b = \omega\left(R_i + \frac{w}{2}\right). \tag{20}$$

3 Results and Discussion

3.1 Bubble Velocity

The effect of channel curvature on bubble characteristics is first examined in Figure 3, where the slip velocity, W, is plotted against β. Numerical simulation results show a clear trend for both $Ca_{TP} = 0.082$ and $Ca_{TP} = 0.0082$. As channel curvature increases, U_B increases and so does W at fixed U_{TP}. It appears that channel curvature affects bubble velocity to a smaller extent at lower Ca_{TP} - this will be discussed in relation to film thicknesses in Section 3.2.

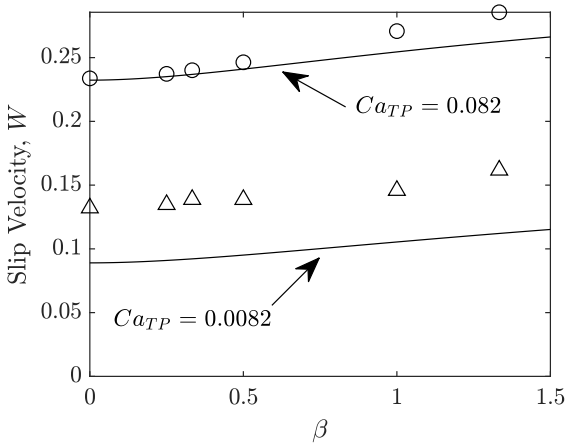


Figure 3: Slip velocity W as a function of channel curvature β. ○: $Ca_{TP} = 0.082$; △: $Ca_{TP} = 0.0082$.

Using Equation 19 a good fit is found between our model and numerical results for $Ca_{TP} = 0.082$ with $\gamma = 3$. The fit is poorer at lower Ca_{TP} is possibly due to a limitation on part of the solver. There is reason to believe that the solver does not work well at lower capillary numbers, as it is unable to obtain the correct values for W even for the straight channel case ($\beta = 0$). Validation with Muradoglu's film thickness expressions detailed in the next section confirm that the values obtained are accurate for the higher capillary number case but not for the lower capillary number case.

It is well known that at low Ca_{TP} and especially at high density ratios, the effect of artificial velocity fields – spurious currents – becomes non-negligible and can have an effect on the bubble velocity [22]. This is because the interFoam solver is uses unstructured meshes, so it is very difficult to implement an accurate scheme to find the second derivative of the indicator function for the calculation of the curvature. This means that the inaccuracy is due to the method, and changing the mesh density would not eliminate this problem. The base grid cell size $\triangle x/w$ at 0.00228 is smaller than that used in other literature [23].

3.2 Film Thicknesses

Film thickness is asymmetrical between the outer and inner bend of the bubble in a curved channel [17]. Figure 4 reflects this trend. As the channel curvature increases, the bubble moves towards the inner bend of the channel, decreasing the inner film thickness $h_{i\infty}$ and increasing the outer film thickness $h_{o\infty}$. As can be seen, our model is in good agreement with $Ca_{TP} = 0.082$ but not so much for $Ca_{TP} = 0.0082$. For the latter, numerical results

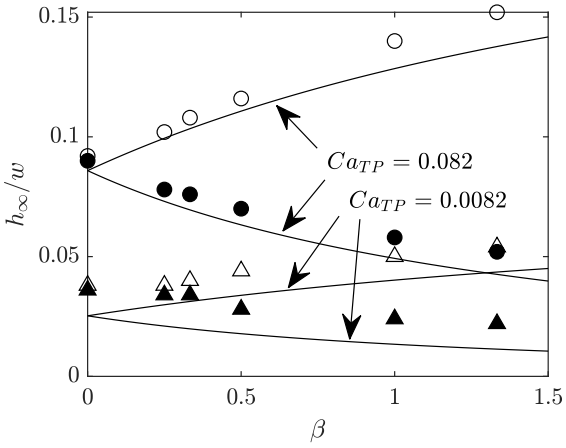


Figure 4: Variation of film thickness with channel curvature. ○: $Ca_{TP} = 0.082$; △: $Ca_{TP} = 0.0082$; filled symbols: $h_{i\infty}$; empty symbols: $h_{o\infty}$. Muradoglu's expression shows a close fit with the model at $Ca = 0.082$.

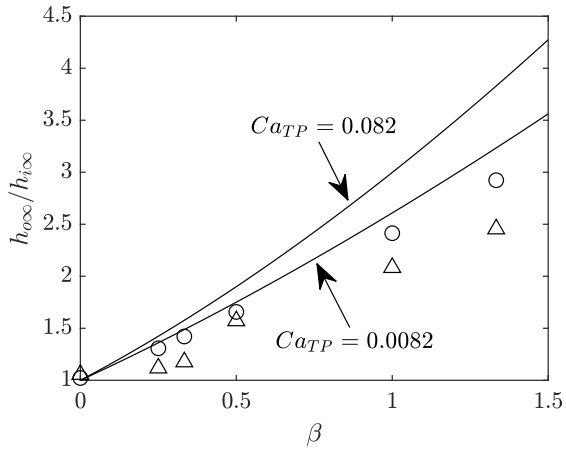


Figure 5: Ratios of the film thicknesses against channel curvature. ○: $Ca_{TP} = 0.082$; △: $Ca_{TP} = 0.0082$.

do not present a smooth trend in terms of increasing $h_{i\infty}$ and decreasing $h_{i\infty}$ with β. Again, this could be attributed to the solver but also in part the mesh density since the absolute difference in film thickness between a curved and straight channel is very small at lower Ca_{TP}.

Validation of our CFD model was done at this stage by comparing film thicknesses with values from Muradoglu's paper [17], with each data point compared individually at the respective bubble capillary number. At $Ca_{TP} = 0.082$, film thickness values fall within 5% of literature values at low β and within 10% at higher β. This deviation goes up 50% at its maximum for $Ca_{TP} = 0.0082$. However, as the trends remain consistent for this value of the capillary number, we can still draw qualitative conclusions from the results.

Once again, from Figure 5, the effect of channel

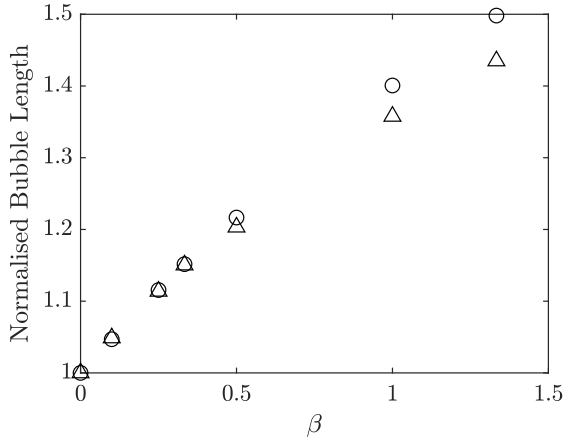


Figure 6: Variation of bubble length with channel curvature. The bubble length is normalised with the length of a bubble with the same volume in a straight channel ○: $Ca_{TP} = 0.082$; △: $Ca_{TP} = 0.0082$.

curvature is more strongly felt for a higher capillary number. The model overestimates the film thickness ratio for both capillary numbers as a result of slightly underestimating both film thicknesses. It might be re-emphasised here that Equation 19 comes from the film thickness as a result of mass balance in the constant film thickness region of the bubble. As such, the slip velocity as a function of β is higher with greater film thickness asymmetry, an observation made by Muradoglu as well [17].

In practical applications, designing for mass and heat transfer should consider the film thickness asymmetry. To improve transfer, it might be more appropriate to position the mass or heat source near the inner bend of the channels, or to design for lower capillary numbers where the film is thinner.

3.3 Bubble Length

Often, images of experimental microfluidic bubbles are captured in 2-D using microscopes [24]. Knowing how to translate from bubble dimensions in 2-D to 3-D volume under different channel dimensions would thus be helpful. Here, we take the bubble length to be the length of the channel center line that marks the sector the bubble lies in (Figure 1). From Figure 6 the bubble appears to elongate by as much as 1.5 times in a curved channel as compared to one of same volume in a straight channel. Approximating a 3-D bubble as a cylinder with spherical caps, the same bubble volume would give a larger surface area for a more elongated bubble. The elongation happens with increasing β, hence changing the channel curvature could be a factor for consideration when designing two-phase flows that require a high surface area-to-volume ratio.

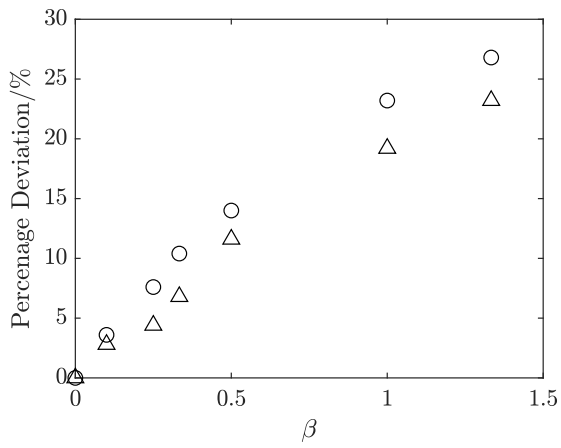


Figure 7: Percentage deviation of bubble head position from channel center line. The percentage deviation is calculated by dividing the distance between the bubble head and the channel center line along the sampling line, by half the channel width. ○: $Ca_{TP} = 0.082$; △: $Ca_{TP} = 0.0082$.

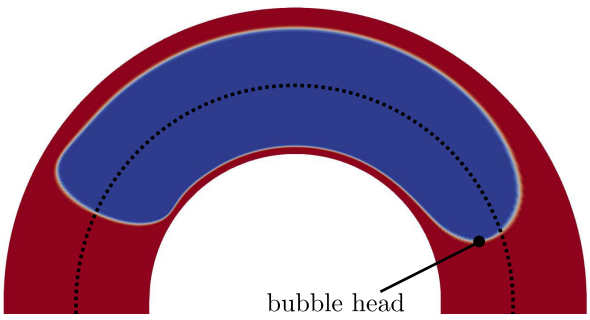


Figure 8: Snapshot of Taylor bubble travelling clockwise in curved channel for $Ca_{TP} = 0.082$. $\alpha = 0$ in the bubble and 1 outside of it.

3.4 Bubble Asymmetry

The final bubble geometry characteristic is the bubble head position relative to the channel center line, something that has not yet been properly characterised in curved channels. The radius of curvature of the bubble head is affected by this asymmetry, and that may play a role in the determination of surface tension properties, and may also determine the transitory film thickness in the same region, which has been shown to correlate with the shear stress experienced by the bubble head [9]. The asymmetry of the nose would also affect the streamlines, and can have an effect of influencing the mixing properties. When travelling in the curved channel, the bubble head is observed to deviate towards the inner bend of the channel (Figure 8) resulting in a smaller radius of curvature and thus potentially higher shear stress. This deviation is quantified in Figure 7 and reaches nearly 30% at high β values. Beyond a certain channel curvature, the bubble was observed to stick to the inner channel wall.

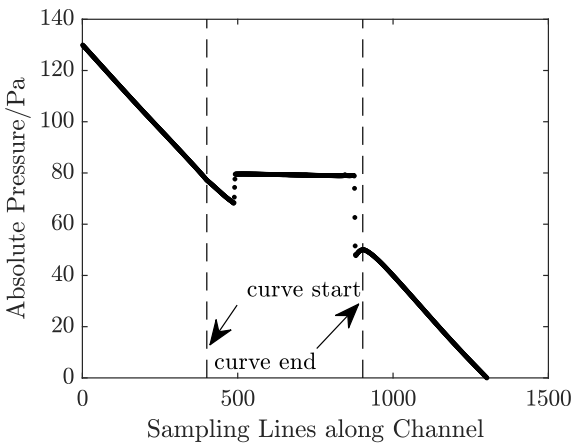


Figure 9: Representative pressure profile along the channel at a certain time step. The dashed lines indicate where the start and end of the curved channel are. $\beta = 1, Ca_{TP} = 0.082$.

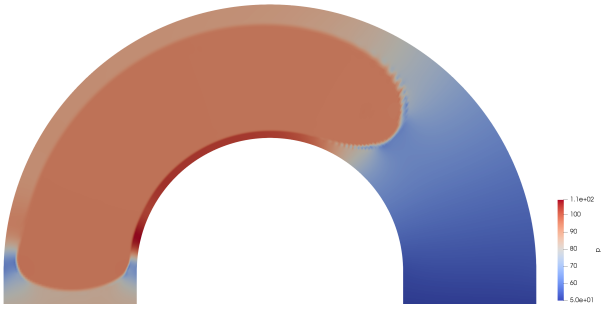


Figure 10: Close-up pressure profile of the bubble in the curved section of the channel (clockwise flow). Red indicates higher pressure, which is in Pa.

3.5 Pressure Drop Across Bubble

We proceed to make use of the bubble head position to define a "bubble center line" which starts from the bubble head and extends through the bubble in an arcline that has the same center coordinates as the channel center line (Figure 1). In order to find the pressure drop across a Taylor bubble, the pressure profile is first obtained through the entire channel. Here we used a domain consisting of a straight channel before and after the same curved channel. The pressure was averaged across the width of the channel at points along the channel center line for regions with $\alpha = 1$ along the channel width, and sampled along the bubble center line otherwise.

Figure 9 shows a representative pressure profile at one time step. Along the straight channels, Poiseuille flow produces a linear pressure drop. At the bubble boundaries, the Laplace pressure due to the surface tension between the fluid and air cre-

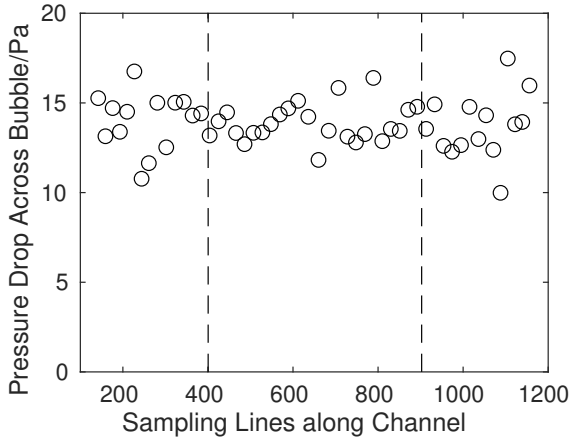


Figure 11: Pressure drop across the bubble as the bubble travels along the channel for $\beta = 1, Ca_{TP} = 0.082$. Similar plots were obtained for other cases. Only a quarter of the data points have been plotted for clarity.

ates a pressure jump, followed by uniform pressure within the bubble. The pressure drop at the bubble head is observed to be larger than that at the bubble tail due to a pressure drop in the liquid film around the bubble [9]. There is a small increase in pressure just ahead of the head of the bubble that can be attributed to the method of sampling.

As seen from Figure 10, the pressure is higher in the fluid region just ahead of the bubble head than at the bubble head itself as a result of high pressures in the liquid films. The averaging of pressure across the width of the channel hence artificially lifts the pressure. Although a detailed analysis is not carried out here, the pressure in the fluid region at the tail of the bubble (Figure 10 is observed to be lower then the bulk fluid around it, which is consistent with Cherukumudi's findings [9].

The pressure drop across the bubble can then be computed as the difference between the two pressure drops at the head and at the tail. Surprisingly, there is no observable trend in pressure drop across the bubble as the shape of the bubble changes with position in the channel (Figure 11). It appears that neither the transition from straight to curved channel nor the presence of a curved channel influences the pressure drop across the bubble.

4 Conclusion

In summary, several important bubble characteristics have been captured in this work. Several aspects of bubble geometry, namely the asymmetric film thicknesses, bubble head position and bubble length. The asymmetry of a Taylor bubble in a curved channel has proven to be an interesting phenomena that can be further investigated more in-depth as it appears to affect a number of other characteristics such as pressure drop and shear stress. The channel geometry in terms of channel curvature was seen to affect Taylor bubbles less significantly at lower capillary numbers.

An expression relating the bubble velocity and the two-phase velocity has also been derived and fits well at capillary number 0.082 for a range of channel curvatures. More simulations across different capillary numbers would be required in order to find better fitting parameters for our model. Finally, the pressure drop has been profiled across the bubble and no trend has been observed for the studied domain, which demonstrates that the bubble geometry has minimal effect on the pressure drop.

5 Future Work

This investigation in 2-D curved geometries is a precursor to more interesting cases as OpenFOAM® makes use of unstructured meshes, meaning that it can easily be extended to more complicated 3-D channels, including periodic geometries such as serpentine or square-wave channels. The methodology introduced in this work to sample the domain by means of slices can be easily adapted to study and quantify the bubble geometry in 3-D. Muradoglu also developed lubrication theory in 3-D to derive an equivalent expression for the film thicknesses that vary in the polar direction [17], which remains to be numerically and experimentally verified.

Finally, the flow patterns in 3-D curved channels are much more complex, as in addition to the effects of pressure drop and centripetal forces, there are dean vortices present that complicate the streamlines of the fluids. Insight into this would be greatly helped by a complete understanding of bubble geometry.

Acknowledgement We would like to thank Dr. Thomas Abadie for his invaluable guidance and support throughout this project.

References

1. Convery, N. *et al.* 30 years of microfluidics. *Micro and Nano Engineering* **2,** 76–91 (2019).
2. Yang, D. *et al. A fast, high throughput, and low-cost microfluidic bioassays for detecting HIV* in *5th Int. Conference on Information Technology and Applications in Biomedicine,*

ITAB 2008 in conjunction with 2nd Int. Symposium and Summer School on Biomedical and Health Engineering, IS3BHE 2008 (2008), 203–206.

3. Deshpande, S. *et al.* Octanol-assisted liposome assembly on chip. *Nature Communications* **7,** 1–9 (2016).
4. Khoshmanesh, K. *et al.* A multi-functional bubble-based microfluidic system. *Scientific Reports* **5** (2015).
5. Ward, K. *et al.* Mixing in microfluidic devices and enhancement methods. *J Micromech Microeng* **25,** 1–33 (2015).
6. Song, H. *et al.* Reactions in Droplets in Microfluidic Channels. *Angewandte Chemie International Edition* **45,** 7336–7356 (2006).
7. Sackmann, E. K. *et al. The present and future role of microfluidics in biomedical research* 2014.
8. Lochovsky, C. *et al.* Bubbles no more: in-plane trapping and removal of bubbles in microfluidic devices. *Lab Chip* **12,** 595–601 (2012).
9. Cherukumudi, A. *et al.* Prediction of the shape and pressure drop of Taylor bubbles in circular tubes. *Microfluidics and Nanofluidics* **19,** 1221–1233 (2015).
10. Klaseboer, E. *et al.* An extended Bretherton model for long Taylor bubbles at moderate capillary numbers. *Physics of Fluids* **26,** 32107 (2014).
11. Aussillous, P. *et al.* Quick deposition of a fluid on the wall of a tube. *Physics of Fluids* **12,** 2367–2371 (2000).
12. Fairbrother, F. *et al.* Studies in electro-endosmosis. Part VI. The "Bubble-tube" method of measurement. *Journal of the Chemical Society (Resumed),* 527–529 (1935).
13. Goldsmith, H. L. *et al.* The movement of single large bubbles in closed vertical tubes. *Journal of Fluid Mechanics* **14,** 42–58 (1962).
14. Bretherton, F. P. The motion of long bubbles in tubes. *Journal of Fluid Mechanics* **10,** 166–188 (1961).
15. Mießner, U. *et al.* An analytical interface shape approximation of microscopic Taylor flows. *Experiments in Fluids* **60,** 75 (2019).
16. Van Steijn, V. *et al.* Predictive model for the size of bubbles and droplets created in microfluidic T-junctions. *Lab on a Chip* **10,** 2513–2518 (2010).
17. Muradoglu, M. *et al.* Motion of large bubbles in curved channels. *Journal of Fluid Mechanics* **570,** 455–466 (2007).
18. Ying, Y. *et al.* Inertial Focusing and Separation of Particles in Similar Curved Channels. *Scientific Reports* **9** (2019).
19. Brackbill, J. U. *et al.* A continuum method for modeling surface tension. *Journal of Computational Physics* **100,** 335–354 (1992).
20. Cifani, P. *et al.* A comparison between the surface compression method and an interface reconstruction method for the VOF approach. *Computers and Fluids* **136,** 421–435 (2016).
21. Aboukhedr, M. *et al.* Simulation of microflow dynamics at low capillary numbers using adaptive interface compression. *Computers and Fluids* **165,** 13–32 (2018).
22. Abadie, T. *Doctorat De L Universit De Toulouse* PhD thesis (Institut National Polytechnique de Toulouse (INP Toulouse), 2013).
23. Hoang, D. A. *et al. Modeling of low-capillary number segmented flows in microchannels using OpenFOAM* in *AIP Conference Proceedings* **1479** (2012), 86–89.
24. Musterd, M. *et al.* Calculating the volume of elongated bubbles and droplets in microchannels from a top view image. *RSC Advances* **5,** 16042–16049 (2015).

Process modelling: different routes to produce FDCA from lignocellulosic biomass

Lanxin FAN
Department of Chemical Engineering, Imperial College London, U.K.

Abstract Large-scale production of biomass-derived chemicals presents a central challenge in green chemistry field. In this paper, a study on different production routes for FDCA (2,5-furandicarboxylic acid), an important intermediate to produce bioplastic, is presented. The work is based on process modelling on Aspen Plus. Two main criteria: Energy saving and profitability help evaluate the performance of each process. Performing heat integration and techno-economic analyses, huge potential has been discovered on ionic liquids and fructose.

Introduction

The energy market's prospect will have a radical change in the following 30 years, as fossil fuel resources are running out, alternative renewable solutions are needed to fill the gap that inevitably will be created. Biomass has a strong potential to be the future resource for production of chemicals in order to replace fossil fuel. Lignocellulosic biomass such as fructose, glucose and cellulose can be transformed into 2,5-Furandicarboxylic acid (FDCA) which is a monomer for the bioplastic. These bioplastics offer a considerably reduced CO_2 emission compared to oil-based plastics. However, nowadays, only a minority of the world annual production of biomass is used to produce bioplastics due to the costly and low efficient process to obtain FDCA from sugars (1).

In the production of FDCA, sugars will be first dehydrated to 5-hydroxymethylfurfural (HMF), which is a key intermediate chemical to obtain various products such as furan-based polymers, fuels and pharmaceuticals (Figure 1). Among these products, 2,5-Furandicarboxylic acid (FDCA) has the most interesting commercial potential and it can be obtained by oxidation of HMF (2) (3). However, several problems are observed in the process. In some cases, the solubility problems limit the use of high concentration reagents, which decreases the productivity and requires more energy for the separation of FDCA. In other cases, purification of the product is technically difficult or energetically costly (3).

Consequently, the objective of this project is to find out the optimal choice for the FDCA production process so that direction for further studies in this area can be forecasted. In order to answer that question, 3 biomass reagents (fructose, glucose and cellulose), 3 suitable solvents (DMSO, water and ionic liquids such as [bmim]Cl) and at least 2 methods of separations will be simulated in the project, using the simulation software: Aspen Plus 8.8. Each combination will be evaluated in term of yield of reaction, purity of product, energy saved with heat integration, utility consumptions, CAPEX, OPEX and environmental impact (mainly the CO_2 emission). Solvent recycling after regeneration will be considered. The optimal option can then be found after comparisons and discussions.

Background

Lots of research has been done at lab scale in the previous work by studying different solvents and catalysts, but the separation of HMF and FDCA have always been a bottleneck. This project aims to summarize the reported work at lab scale and repeat them at an industrial scale.

It is reported that the conversion of sugars to HMF and HMF to FDCA are higher when using separate solvent systems (4) (5). Currently few studies are reported for the direct synthesis of FDCA from sugars. However, it is still an interesting idea to produce FDCA directly from sugars because this permits to avoid separation of HMF which is very tedious (6). Thus, this represents a more sustainable and economic way. Some solvents which were

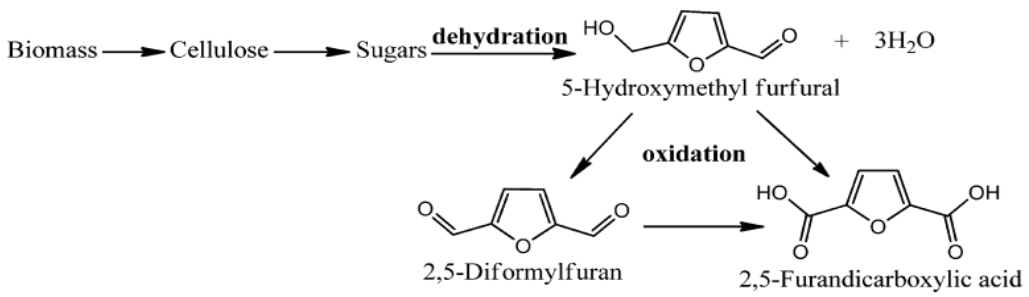


Figure 1. Conversion of biomass sugars to Furan Compounds

Compound	Abbreviation	Formula	Molecular weight	Boing point (°C)	Solubility
Fructose	-		180.1	>200	Soluble in water, DMSO and [bmim]Cl
Glucose	-		180.1	>200	Soluble in DMSO and [bmim]Cl
Cellulose	-		180.1	>200	Soluble in [bmim]Cl
5-hydroxymethylfurfural	HMF	$C_6H_6O_3$	126.1	291	Soluble in water, DMSO and [bmim]Cl, higher affinity with MIBK
2,5-furandicarboxylic acid	FDCA	$C_6H_4O_5$	156	420	Low solubility in GVL, 1g/L in water
Dimethyl sulfoxide	DMSO		78.1	189	-
Methylisobutylketone	MIBK		100.1	117	-
γ-valerolactone	GVL	$C_5H_8O_2$	100.1	207	Miscible with water
1-butyl,3-methyl imidazolium chloride	[bmim]Cl		175,6	>200	-

Table 1. Details and properties of main chemicals present in the system (7) (8)

mentioned in previous work such as DMA, DMF, THF, acetonitrile are no longer suitable in our case, because they are unable to carry out the oxidation of HMF in further processing and they are not considered as green solvents for their toxicity and volatility. Consequently, only three widely-studied solvents are considered in this project—Water, DMSO, [bmim]Cl (9).

Many studies have been carried out for the dehydration of sugars to HMF. During the last 30 years of studies in this area, the acid-catalyzed triple dehydration of fructose to HMF is regarded as the most efficient and convenient synthetic route. It has been demonstrated that almost full conversion of fructose and 100% of HMF selectivity can be achieved (10). Glucose and cellulose can also be used as feedstock but the choice of the solvent is crucial to obtain high yields. Fructose is the most compatible reagent to produce HMF since it can react in water, DMSO and ionic liquids (11). Glucose can be utilised in DMSO and ionic liquids while cellulose can only be chosen when using ionic liquids (12) (7) (13) (14) (15). The price, yield of HMF and loading are different for fructose, glucose and cellulose, as shown in the table below.

	Fructose	Glucose	Cellulose
Loading	30%	10%	10%
Yield of HMF	99%	70%	60%
Price	453	360	350

Table 2. Loading, yield of HMF and price for different sugars (16) (12) (13) (14) (17) (18) (19)

The yield of HMF when using glucose and cellulose is lower, only 70% and 60% respectively. The generation of side-products, which is mainly humin polymers, should be precipitated by adding water before the second reaction (7).

Solvent is added at 100 kg/h . Water is a cheap green solvent; thus, recycling is not necessary. However, a large amount of second liquid (MIBK) must be added to form a biphasic system during the first reaction so that the HMF produced can be extracted. This prevents HMF from reacting with fructose to produce side-
products. The use of organic solvent (MIBK) induces problems of toxicity and volatility. As a result, the recycling of MIBK is indispensable (9). When using DMSO, and ionic liquids as solvent, due to the expensive price, solvent is recommended to be recycled.

Oxygen and water should be added in the oxidation reaction. We suppose that oxidation needs 20%wt of water for DMSO and ionic liquid (20) and oxygen is 20% in excess. The temperature and pressure conditions for the two reactions have all been reported in the literatures. The most representative ones are reported in Table 3.

	Dehydration	Oxidation
Water + MIBK	150°C, 10 bar	120°C, 10 bar
Water + GVL	180°C, 20 bar	100°C, 40 bar
DMSO	120°C, 3 bar	120°C, 10 bar
[bmim]Cl	100°C(F)/120°C(G)/ 150°C(C), 3 bar	120°C, 10 bar

Table 3. Operating conditions for dehydration and oxidation reaction (21) (22)

The separation of FDCA can be either thermal or physical, two methods were proposed in literature: solvent removal by distillation and crystallization (23). However, the use of distillation column is not accurate. Solvent such as MIBK and DMSO require high energy to be boiled due to their high boiling point. A more sophistic method is to use multiple effect evaporators, the heat generated from the previous evaporator's condenser is used to provide heat for the next evaporator (24). The crystallisation is another common method. Solvents such as water or acid (HCl) are added in the crystallizer to precipitate the FDCA (9) (7).

Regeneration is a vital part in the sustainable industrial production of FDCA, it is needed for the solvent recycling and it helps control the impurities amount. The regeneration can be done either with a

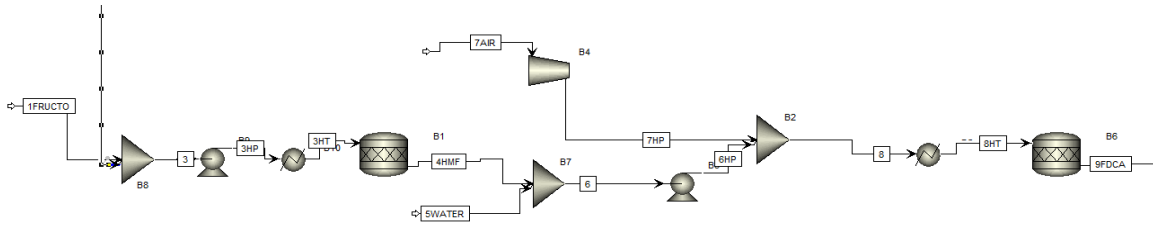


Figure 2. Flowsheet of FDCA production section with fructose as reagent

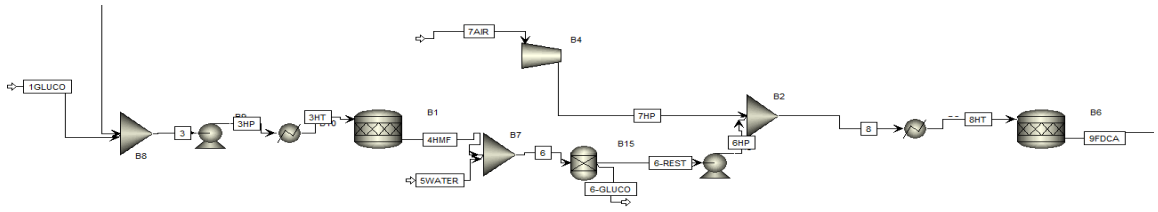


Figure 3. Flowsheet of FDCA production section with glucose or cellulose as reagent

distillation column or a simple flash depending on the water content requirement in the input flow. DMSO, MIBK are allowed to have a very low water content (less than 1%wt) while [bmim]Cl can have up to 10%wt water in the recycling loop (21). GVL should be added with water at a ratio of 50:50wt (21). Once the solvent regeneration is done, a purge is implemented to prevent the accumulation of impurities in the system. Due to the solvent loss, a solvent supplement is needed to maintain a constant flowrate.

Methods

Background research is a great entry point to launch the project, by exploring previous experimental data on reaction conditions, possible synthesis routes, thermodynamic models and parameters etc., it is easier to gain knowledge about this area. Process modelling is the fundamental work of this project, the Aspen Plus Version 8.8 process simulator was chosen to simulate the dehydration, oxidation, separation of FDCA and regeneration of solvent whose conditions have been described in the previous chapter. The data obtained from the process modelling can be used to perform heat integration and to calculate utilities. With the energy and material input, economic analyses could be carried out regarding the capital cost (CAPEX), operating cost (OPEX) and payback period. The main criteria for choosing the best production route of FDCA is through the comparison of energy requirement and economic cost. Further work can be done on the Life Cycle Assessment (LCA) in order to estimate the environmental impact of the process in term of CO2 emission.

1) General description of the process

The upstream reaction section has almost the same configuration for all the cases. The dehydration of lignocellulosic biomass to HMF and the oxidation of HMF to FDCA are carried out separately in two reactors. They are both using the model stoichiometric reactor (RStoic) since the reaction kinetics are unknown but stoichiometry and yield of the reaction are known from the literature. Mixers are used to add input streams such as sugars, solvent, water and oxygen to the system, which enter at ambient temperature and pressure (25°C and 1 bar). Pumps, compressors and heat exchangers are implemented to adjust the flows to reach the operating reaction conditions in the reactor. A pump is used to increase the pressure of liquid flow while air or pure oxygen is pressurized with an isentropic compressor. Their efficiency is estimated at 90% (23) and the discharge pressure was mentioned previously in Table 3. A heater then raises the temperature to the required value (Table 3). There is a slight difference in the flowsheet according to the use of reagent. While using glucose and cellulose as reagents, the yield of HMF is only 60-70%. Consequently, a separator is implemented between the dehydration and oxidation reactor to simulate the precipitation of side products.

After the conversion of FDCA, the system contains the product FDCA, some unreacted reagents, side products, a large amount of solvent and oxygen in excess. The flow is introduced directly in a flash (at 100°C and 10 bar or adiabatic). This first step of separation aims to separate the liquid phase from gaseous phase. Then, two methods cited previously can be chosen in order to separate FDCA from other chemicals. The purity of FDCA for each flowsheet is set to be around 99%. Multiple effect evaporation is suitable for separating FDCA from DMSO. Three evaporators are in series to

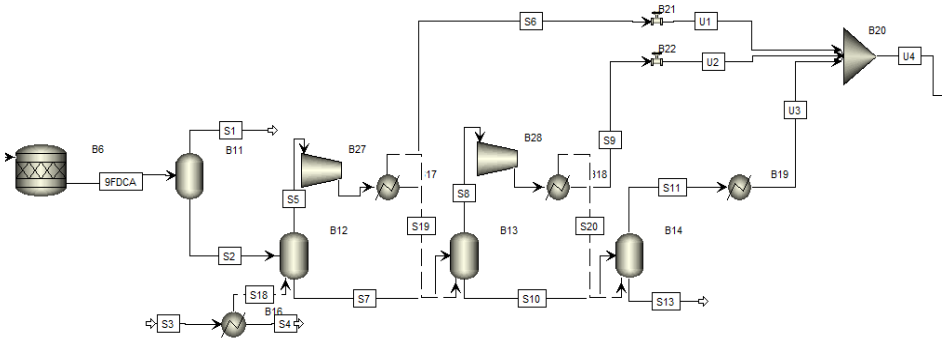


Figure 4. Flowsheet of FDCA separation section using multiple effect evaporators

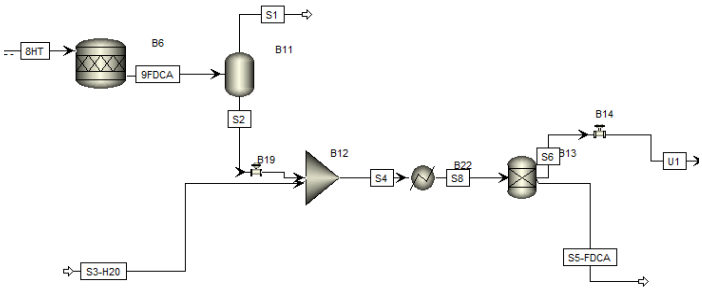


Figure 5. Flowsheet of FDCA separation section using crystallizer

achieve the required purity. The ranges of pressions in each evaporator is 0.5 bar, 0.1 bar, 0.01 bar respectively which are carefully selected to ensure a better separation. At the top of each evaporator, gaseous DMSO is collected since its boiling point is lower than FDCA. The stream is introduced in a compressor then in a condenser, where the vapor phase is completely condensed to recover energy for the next evaporator. The compressor aims to increase the stream's pressure so that the outlet boiling point of the condenser is higher than the temperature in the next evaporator with a minimum temperature approach of 5 °C. The necessary heat for the first evaporator is provided by a condenser where a saturated steam at 1 bar and 120 °C passes through. Its flowrate is 18,9 kg/h and 17.4 kg/h for the case of using fructose and glucose respectively in order to achieve the required purity. These values were determined with the help of design specification by setting the purity of FDCA at 99%. The steam could be generated with the heat released from the reaction section. The heat integration will be discussed in below section 2).

Another widely-used method for the separation of FDCA is crystallisation. The crystalliser is simulated with the separator model where a purity of 100% and a recovery rate of 99% are fixed. A valve and a heat exchanger help to condition the outlet stream to the required pressure and temperature in the crystalliser, which is 25 °C and 1 bar. Water is added in the system with the help of a mixer while using DMSO and [bmim]Cl as solvent (9) (7). Hydrochloric acid (HCl) is added when water is used as solvent (25). The necessary amount is calculated according to the literature and the values are below.

Solvent system	Reagent	Amount of solvent added for crystallisation (kg/h)
Water + MIBK	Fructose	0.32 (HCl)
Water + GVL	Fructose	-
DMSO (9)	Fructose	4.51
	Glucose	59.16
[bmim]Cl (7)	Fructose	0
	Glucose	24.35
	Cellulose	25.60

Table 4. Amount of solvent (water or HCl) added for crystallization. Flowrate of HCl on the first line is presented for the case of water+MIBK, for all the other cases, water helps precipitate FDCA

As for the recycling section. The choice of regeneration method depends on the boiling point of the chemicals and the product specification. Unlike the reaction and separation section, which is equivalent for all the scenarios, the recycling section has 4 different cases and they are detailed below:

- Regeneration of MIBK from HMF

The outlet stream of the dehydration reactor forms a biphasic system, where water and MIBK are two immiscible solution, and HMF, as the solute, is dissolved in the MIBK. In order to recycle the MIBK at a very low water content (less than 1%) and to separate the HMF for the oxidation reaction, multiple effect evaporators are implemented. HMF remains at the bottom as the heavy key while MIBK is collected at the top of each evaporator. Since the

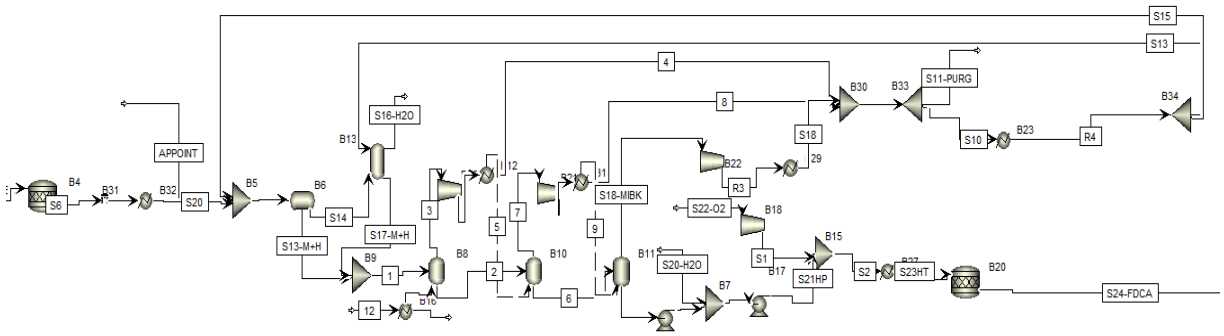


Figure 6.Flowsheet for regeneration and recycling of MIBK from HMF

boiling point of HMF is quite high (291 °C), it is not recommended to use a simple distillation column. Heat integration consideration also presents in this scenario. The recycled MIBK is collected with a mixer and a purge of 1% is included to avoid inert accumulation in the system. The stream is then cooled down to 25 °C, ready to be recirculated. A splitter helps distribute the recycled MIBK to the decanter and to the extraction column which operate both at 25 °C and 1 bar. A MIBK supplement of 11 kg/h at 25 °C and 1 bar is connected to the mixer in order to maintain the constant solvent flowrate.

- Regeneration of DMSO from water

The water content requirement in the DMSO is very strict (less than 1%). Thus, regeneration can either be multiple effect evaporation or distillation which is a cheaper choice. Since the boiling point of water is moderate (100 °C), it is possible to implement a simple vacuum distillation column. The 8-stage column is simulated with RadFrac model at 0.01 bar with a feed at third stage, which allows a 99.9% recovery rate of DMSO. A purge of 1% (simulated with a splitter) and a DMSO supplement (simulated with a mixer) are implemented. The flowrate is 1 and 2.5 kg/h for crystallisation and multiple effect evaporation respectively. A pump is used to adjust the outlet stream of distillation column to ambient pressure. The stream is then recirculated and mixed with the feed stream.

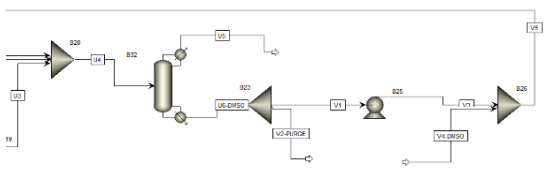


Figure 7. Flowsheet for regeneration and recycling of DMSO from water

- Regeneration of [bmim]Cl from water

The water content requirement is loose in the [bmim]Cl recycling loop, up to 10%wt of water (10kg/h) content can be accepted and has little effect on the reaction. As a result, it is recommended to select a simpler and cheaper way of separation. According to phase diagram generated from Aspen+, a simple flash can separate [bmim]Cl from water and other impurities. The evaporation is supposed to be taken place at 115 °C, considering the minimum temperature approach for heat integration. A design specification and a balance variable were included in flash block so that the water flowrate can be no more than 10kg/h by updating the flash's pressure. The value is calculated to be 0.045 bar.

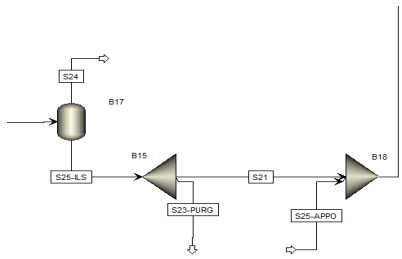


Figure 8. Flowsheet for regeneration and recycling of [bmim]Cl from water

- Recycling of GVL and water

The solvent is fed at 100 kg/h with a composition of 50%wt water: 50%wt GVL, the fructose loading is 15% and the yield of HMF is 70%. FDCA is separated with crystallisation in block B14. A flash is used to recycle all the GVL and only 50 kg/h of water. No solvent supplement is needed as there is no solvent loss. A design specification was added in the block flash, which helps to fix the water flowrate at 50 kg/h. The flash's pressure was set at 0.069 bar through the design specification while its temperature is fixed at 40 °C.

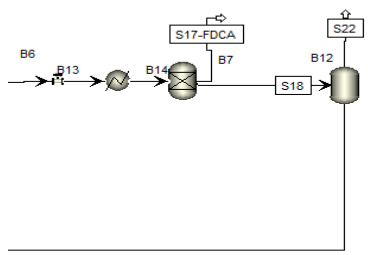


Figure 9.Flowsheet for regeneration and recycling of GVL and water

Solvent	Reagent (Loading, yield)	Separation	Recycling
Water (100kg/h) + MIBK (700kg/h)	Fructose (30%, 0.99)	Crystallization (25°C, 1bar)	MIBK-Multiple effect Evaporators
Water (50kg/h) + GVL (50kg/h)	Fructose (15%, 0.7)	Crystallization	Water + GVL-Flash
DMSO (100kg/h)	Fructose (30%, 0.99)	Crystallization	DMSO-Distillation
		Multiple effect Evaporators	DMSO-Distillation
	Glucose (10%, 0.7)	Crystallization	DMSO-Distillation
		Multiple effect Evaporators	DMSO-Distillation
[bmim]Cl (100kg/h)	Fructose (30%, 0.99)	Crystallization	[bmim]Cl -Flash
	Glucose (10%, 0.7)	Crystallization	[bmim]Cl -Flash
	Cellulose (10%, 0.6)	Crystallization	[bmim]Cl -Flash

Table 5. Summary of process unit of each flowchart

1) Thermodynamic model

The selection of the thermodynamic model among the available options in Aspens+ simulator is done according to previous studies. COSMOSAC is the most suitable thermodynamic model for ionic liquids (26). The simulator allows to directly introduce the properties of ionic liquids from its database. However, properties parameters of others components such as oxygen, fructose, water, HMF and FDCA should be added manually in the model. The Electrolyte NRTL model is used for other cases (23). The estimation of components' properties is introduced directly from its database.

2) Heat integration

The first degree of heat integration is to use the energy released from the second reaction to provide heat needed in the separation and regeneration section. The first reaction is exothermic in most of the cases of study. However, the generated heat is considered negligible compared to the energy released by the second reaction. The only exception for the heat is when using COSMOSAC as thermodynamic model, this is considered as a problem with the model to be fixed in future study.

The second part of heat integration consists to construct an exchanger network. The heat exchangers before decantation, crystallization or recycling helps to decrease the temperature to operating condition, which is normally 25 °C, the heat recovered can be provided to heaters before the reactors. These heaters help heat up the streams to required reaction conditions. However, some heat exchangers before the reactors are condensers due to the high-temperature streams brought by the recycling loop or compressor. The construction of exchanger network is performed with the help of Aspen Energy Analyser. The inlet and outlet temperatures, mass flowrate and heat capacity of each flow were entered manually in the software, the pinch points for hot stream and cold stream are then calculated. Different exchangers can be connected below and above the pinch points by us. The saved energy, cold utility, hot utility and cost are available.

3) Economic analyses

The economic analysis is carried out in two part. Variable costs and benefit are calculated depending on the price of raw materials per unit and feedstocks. Table 6 shows the current prices for most used chemicals and product FDCA in the system. In addition to chemicals, utilities (cooling water, steam and electricity) is another major cost which requires constant input. The cooling water is available at 20 °C to satisfy a minimum temperature approach of 5 °C since most of the cases need to be cooled down to 25 °C. The temperature increase should not be higher than 15 °C. As for the choice of steam, several pressures are available with different costs (Table 7). For the same reason of minimum temperature approach, the saturated steam at 186 °C is selected and its cost is 7.5 $/metric ton. The price for cooling water and electricity is 0.00812 $/metric ton and 3.81 $/kW respectively (27) . The amount of cooling water and steam needed in the system is calculated with the results of heat integration. Once the cold utility and hot utility are determined. the mass flowrate can be determined with the heat capacity of water, heat of vaporisation of steam and temperature rise. Multiplying by the given price, the utility costs after heat integration can be obtained. Thus, the variable cost which is called as well OPEX (operating cost) is composed of raw material cost, solvent cost, oxygen and process water cost and utility cost. The man power and maintenance costs are not being considered here. The benefit of selling FDCA minus the variable costs gives a net profit of the process per year.

	Fructose	Glucose	Cellulose	O2	FDCA
USD/t	453	360	350	250	1900
	MIBK	GVL	[bmim]Cl	DMSO	Water (25 °C)
USD/t	850	3000	7851	1489	0.9

Table 6. Prices for used chemicals in $/Metric ton (28) (29) (30) (31)

P (bar)	Tsat (°C)	ΔHvap (kJ/kg)	Cost (USD/t)
1	120	2205.281	5.03
3	147	2126.197	6.17
10	**186**	**1995.01**	**7.50**
17	207	1913.6	8.20
41	253	1659.369	9.96

Table 7. Utility steam data: Saturation conditions, Enthalpy of vaporisation and cost

The second part consists a study of profitability. The capital cost estimation considers both equipment and associated plant bulks as design-based installation models. The results are directly provided with Aspen Plus Economics. The same costing engine helps compare the relative costs of one design versus another. Considering the capital cost (CAPEX), the payback period and rate of return can be calculated to help compare the profitability of each process. The payback period refers to the time required to recoup the funds expended in an investment (CAPEX) and to reach the break-even point. Shorter the payback period is better the process is. The rate of return (RoR) is the net gain or loss on an investment over a specified time period, expressed as a percentage of annually net profit out of annually variable cost.

Results

Table 5 summarizes the feedstocks and process units of each flowsheet in the process modelling work. At this basis, the heat integration shows us results in the following tables. Considering the oxidation reaction, separation and regeneration, process 2, 3, 6, 8 are self-sufficient. Moreover, the energy released from these processes can be used to generate electricity to provide to compressors.

	Q (kJ/h)	Q oxydation	Q evap	Q reg	Total heat
1	Water_MIBK_F_3evap	-139234	168064	0	28830
2	Water_MIBK_F_4evap	-118751	82909	0	-35843
3	Water_GVL_F	-45372	0	14951	-30421
4	DMSO_F_Crysta	-117733	0	118860	1128
5	DMSO_G_Crysta	-32281	0	246313	214032
6	DMSO_F_Evap	-129214	43492	51127	-34595
7	DMSO_G_Evap	-30461	40057	49064	58660
8	BMIMCL_F_Crysta	-135018	0	90127	-44890
9	BMIMCL_G_Crysta	-31741	0	124580	92839
10	BMIMCL_C_Crysta	-27213	0	127189	99976

Table 8. Heat integration between oxidation reaction, separation section (the first evaporator in multiple effect evaporation), regeneration section (distillation column or flash). Negative sign stands for releasing energy and positive sign stands for the needed energy. In the column for total heat: the red bars mean that the process is releasing more energy than needed, and the blues bars mean the opposite. Process no.2 is when using water as solvent, MIBK as second solvent, fructose as reagent and 4 evaporators in multiple effect evaporation. Process no.3 is when using water as solvent, GVL as second solvent, fructose as reagent. Process no.6 is when using DMSO as solvent, fructose as reagent and multiple effect evaporation as separation method. Process no. 8 is when using [bmim]Cl as solvent, fructose as reagent and crystallization as separation method.

Table 9 shows the heat generated and needed by condensers and heaters in each scenario. The saved energy, cold and hot utility and CAPEX are calculated by Aspen Energy Analyzer.

	Q (kJ/h)	E1	E2	E_decanter	E_recycling	E_crysta	Heat saved	Cold utility	Hot utility	CAPEX
1	Water_MIBK_F_3evap	64385	5047	-66340	-236017		66857	235500	2575	43558
2	Water_MIBK_F_4evap	64385	6771	-66340	-220314		43572	243083	27585	24125
3	Water_GVL_F	52564	-33038			-26209	33038	26209	19526	11602
4	DMSO_F_Crysta	16427	8422			-30498	9458	21040	15391	12360
5	DMSO_G_Crysta	12897	7752			-25711	7214	18498	13435	12083
6	DMSO_F_Evap	26397	9982		-61960		22093	39867	14286	12646
7	DMSO_G_Evap	14065	7943		-50762		11295	39468	10714	11591
8	BMIMCL_F_Crysta	-3246	5840			-28167	0	31413	5840	0
9	BMIMCL_G_Crysta	366	-982			-22206	0	23189	366	0
10	BMIMCL_C_Crysta	365	-943			-22015	0	22958	365	0

Table 9. Table of heat integration between condensers and heaters. In green shows the heaters and in orange the condensers. The results of saved heat, cold and hot utility and CAPEX are shown in bar chart proportionally. Process 1 and 2 (Water as solvent, MIBK as second solvent) have a huge energy save but the utility and cost are high. Process 8-9 using [bmim]Cl as solvent and the heat integration is not feasible. However, the utility requirement is the lowest.

Figure 10 shows the cold and hot utility consumption needed in the system. They are calculated by summing up the results of the two heat integrations.

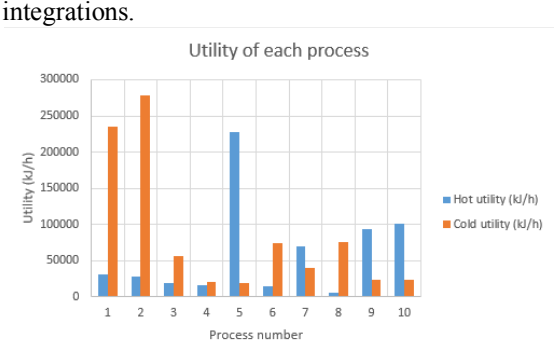


Figure 10. Hot and cold utility needed in each process in kJ/h. The orange bars refer to cold utility while the blue bars stand for cold utility. Process no.1 (Water+MIBK+Fructose+3 evaporators) , Process no.2 (Water+MIBK+Fructose+4 evaporators) , Process no.3 (Water+GLV), Process no.4 (DMSO+Fructose+Crystallisation), Process no.5 (DMSO+Glucose+Crystallisation), Process no.6 (DMSO+Fructose+Multiple Effect Evaporation), Process no.7 (DMSO+ Glucose + Multiple Effect Evaporation), Process no.8 ([bmim]Cl+Fructose), Process no.9 (bmim]Cl+Glucose), Process no.10 (bmim]Cl+Cellulose)

From the utility consumption, the estimation of utility costs is presented in detail in Table 10 for more information. The other variable costs and the benefit of FDCA selling are estimated from the prices found on the Internet (28) (29) (30) (31) and the flowrate given by Aspen Plus. (The detailed Table 11 is given in the Supplementary Information section). The Figure 11 shows the net benefit of each scenario.

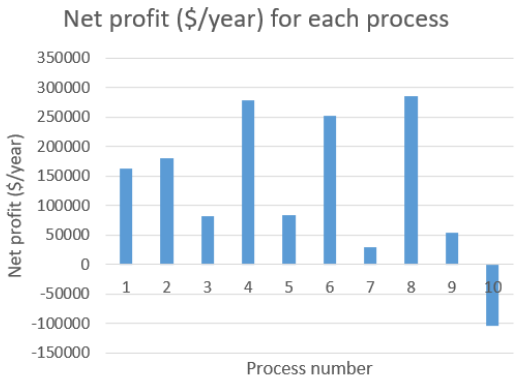


Figure 11. Net profit ($/year) of each scenario. Net profit per annum is calculated by subtracting annually variable costs from total benefit per year. The numbers of scenarios are the same as described in Figure 10.

The capital cost is given directly by Aspen Plus Economics as shown in Figure 12. For accurate values of CAPEX, please inquire in Table 12.

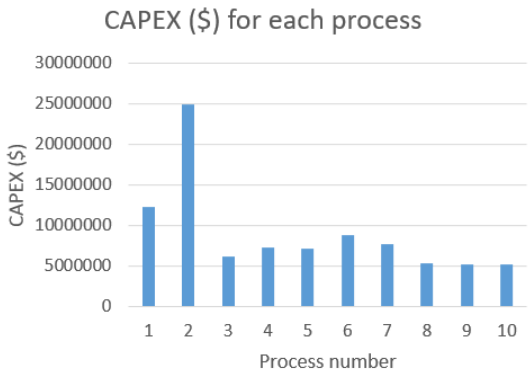


Figure 12. Capital Cost (CAPEX) of each scenario in $. The numbers of scenarios are the same as described in Figure 10.

The payback period and rate of return of each scenario are shown in Figure 13 and Figure 14 respectively. The accurate values are presented in Table 12 in Supplementary Information section.

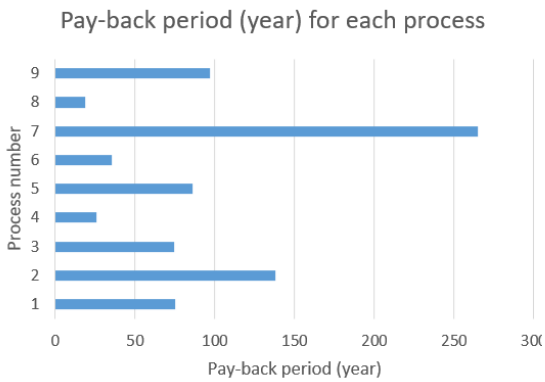


Figure 13. The payback period (year) of each scenario, supposing the inflation rate is 0. It refers to the amount of year to reach the break-even point. The numbers of scenarios are the same as described in Figure 10.

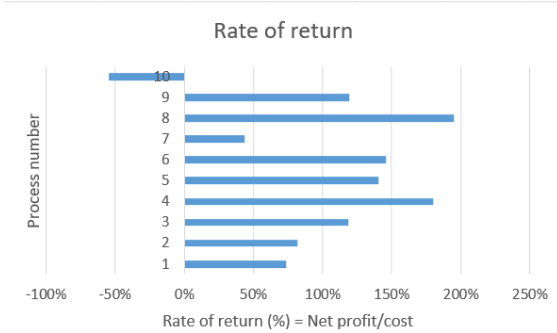


Figure 14. The rate of return of each scenario. The numbers of scenarios are the same as described in Figure 10.

Discussion

The utility costs have a strong connection with both the economic aspect and environmental impact. A smaller utility cost means greater benefit and lower CO2 emission. The heat integration is an efficient way to improve the process's performance and reduce the utilities. However, more energy saved does not mean less utility costs. The determined criteria should be the total hot and cold utilities needed in the process instead of the amount of energy saved. Even though the process no.1 and no.2 save more energy by constructing an exchanger network, it requires more energy input than the other processes. The hot utility is supposed to be saturated steam at 186 °C and cold utility is cooling water at 20 °C. The water is heated up and turns into steam which can be in turns used to provide electricity in a turbine.

While all the other conditions remain the same, it can be concluded that using glucose as reagent requires always more energy input than using fructose, and cellulose is the most energetically demanded. That is due to the fact that loading of glucose and cellulose are lower than that of fructose, the productivity of FDCA and its concentration in solvent are smaller. Thus, the separation of FDCA will be more difficult. It is noticeable that the processes using MIBK require a large amount of cold utility. That is because flows constantly need to be cooled down from a very high temperature (around 200 °C) for decantation or recycling. The three lowest-energy consumption scenarios are process no.3 (DMSO + Fructose + Crystallisation), process no.4 (Water + GVL) and process no.7 ([bmim]Cl + Fructose) by descending order. Surprisingly, even though 3 evaporators and compressors are implemented in the process no.6 (DMSO + Fructose + Multiple effect evaporation), it still has a low hot utility requirement and an average cold utility demand, which make it an interesting case.

In terms of economic, the same conclusion is reached by comparing the three reagents. Processes using fructose (process no. 4, 6, 8) have a greater

profit than processes using glucose (process no.5, 7, 9), and far better than using cellulose (process no.10), for which the net profit is negative. In terms of solvent, even though [bmim]Cl is the most expensive solvent, it is the best choice to maximise the profit, which is slighter better than using DMSO, and far greater than using water. While using DMSO as solvent and fructose or glucose as reagent, by comparing the process no. 4 and no.6, process no.5 and no.7, we can conclude that crystallisation as separation method is slightly cheaper than multiple effect evaporation.

The CAPEX may be over-evaluated by Aspen Plus Economics because the utility costs given by aspen are 10 times higher than the actual ones. However, the ratio between different process is correct as the same costing estimation engine is used throughout the modelling. The process no.1 and no.2 which use MIBK as solvent, need an extremely high investment due to the complexity of process, Except for these two scenarios, the other processes have the same range of CAPEX which is between 5 million and 9 million. Implementation of multiple effect evaporators are more expensive than a simple crystallizer.

In term of payback period and rate of return the process 8, 4, 6 are the best by descending order. These three most profitable processes require 19, 26, 35 year respectively to reach the break-even point. However, since the real CAPEX may be lower, it might take only 6-12 years, which is a normal waiting period to start making profits. The rate of return explains the profitability of process from another aspect. Taking process no.8 ([bmim]Cl + fructose) as example, 286 thousand dollars are gained while spending 146 thousand dollars, 200 % of return rate means the FDCA production process is efficient and highly profitable.

Conclusions

Techno-economic analyses have shown a great potential market for biomass-derived commodity chemicals and a possibility for industrial scale-up. Most of the cases in this project show an attracting rate of return and payback period with consideration of environmental impact. Combining the energy saving criteria with the profit driving selection. The most potential production route uses [bmim]Cl and fructose, which are both the best solvent and reagent among our choices. The second-best process is the scenario using DMSO as solvent, fructose as reagent and crystallisation as separation method. Its overall performance is just slightly worse than the "[bmim]Cl + fructose" production route.

More other work could be done to improve the credibility of this project. The production costs may be lower if chemicals are produced at larger scales in continuous flow. Thus, an optimisation should be performed in order to find the optimal operating flowrate so that the largest benefit can be reached. A thorough techno-economic analyses of the two best production routes should be performed, especially an estimation of CAPEX. Life Cycle Assessment is another important method to evaluate the environmental impact.

References

1. E4tech, Re-Cord, Wur. From the Sugar Platform to biofuels and biochemicals. s.l. : Directorate-General Energy, 2015. N° ENER/C2/423-2012/SI2.673791.

2. Replacing fossil based PET with biobased PEF; Process analysis, energy and GHG balance. Eerhart, A. J.J.E.; Faaij, A. P.C.;Patel, M. K. 4, s.l. : Energy & Environmental Science, 2012.

3. Rapid, High-Yield Fructose Dehydration to 5-Hydroxymethylfurfural in Mixtures of Water and the Noncoordinating Ionic Liquid [bmim][OTf]. Al Ghatta, Amir; Wilton-Ely, James D.E.T.; Hallett, Jason P. 19, s.l. : ChemSUSChem, 2019, Vol. 12, pp. 4452-4460.

4. Selective Synthesis of 2,5-Diformylfuran and 2,5-Furandicarboxylic Acid from 5-Hydroxymethylfurfural and Fructose Catalyzed by Magnetically Separable Catalysts. Zhenzhen Y., Wei Q., Rongxin S., Zhimin H. 1, s.l. : Energy & Fuels, 2016, Vol. 31.

5. Highly Efficient Oxidation of 5-Hydroxymethylfurfural to 2,5-Furandicarboxylic Acid with Heteropoly Acids and Ionic Liquids. Ruru C., Jiayu X.,Dongxia Y.,Huixian D.,Xingmei L., Suojiang Z. 12, s.l. : ChemSUSChem, 2019, Vol. 12, pp. 2715-2724.

6. Using Sub/Supercritical CO2 as "Phase Separation Switch" for the Efficient Production of 5-Hydroxymethylfurfural from Fructose in an Ionic Liquid/Organic Biphasic System. Chunyan S.,Jiayu X., Xiaomin L., Xingmei L., Suojiang Z. 2, s.l. : ACS Sustainable Chemistry & Engineering, 2016, Vol. 4, pp. 557-563.

7. Strategies for the Separation of the Furanic Compounds HMF, DFF, FFCA, and FDCA from Ionic Liquids. Amir A. G., James D. E. T. W. , Jason P. H. 7, 2019, ACS Sustainable Chem. Eng., p. 16483−16492.

8. Siddarth H. K., Kefeng H., Kevin J. B. Oxygenated commodity chemicals from chemo-catalytic conversion of biomass derived heterocycles. Department of Chemical and Biological Engineering, University of Wisconsin-Madison. Madison : s.n., 2018.

9. Strategic Designing on Selection of Solvent Systems for Conversion of Biomass Sugars to Furan Derivatives and Their Separation. Sampath, G.,

Kanagaraj, I. A., Sowthri R.; Kannan, S. 2017, ACS Sustainable Chem. Eng, p. 5373−5381.
10. Acid-Catalyzed Dehydration of Fructose to 5-(Hydroxymethyl)furfural. Yunxiang, Q., Nils, T., Zhenshan, H. 2, 2015, Recycl. Catal., pp. 36-60.
11. 5-Hydroxymethylfurfural (HMF). A Review Focussing on its Manufacture. Kuster, Dr. B. F. M. 8, s.l. : Starch, 1990, Vol. 42, pp. 314-321.
12. Conversion of glucose and cellulose into value-added products in water and ionic liquids. Jinliang S., Honglei F., Jun M. and al. 10, s.l. : Green Chemistry, 2013.
13. The Highly Selective and Near-Quantitative Conversion of Glucose to 5-Hydroxymethylfurfural Using Ionic Liquids. Sanan E.,Agnieszka B.,James D. E. T. Wilton-Ely ,Jason P. H. s.l. : PLOS, 2016.
14. Direct Catalytic Conversion of Cellulose to 5-Hydroxymethylfurfural Using Ionic Liquids. Sanan E., Paraskevi F., James D. E. T. Wilton-Ely. 32, s.l. : Inorganics, 2016, Vol. 4.
15. Catalytic dehydration of hexose sugars to 5-hydroxymethylfural. Xianli Z., Chang Z., Qian W., Gang Y. 1, s.l. : Biofpr, 2019, Vol. 13, pp. 153-173.
16. Production of 2,5-furandicarboxylic acid (FDCA) from starch, glucose, or high-fructose corn syrup: techno-economic analysis. Luana D., Sadra S. and al. 5, s.l. : Biofpr, 2019, Vol. 13, pp. 1234-1245.
17. Phase Modifiers Promote Efficient Production of Hydroxymethylfurfural from Fructose . Román-Leshkov Y., Chheda J.N., Dumesic J.A. s.l. : Science, 2006.
18. Fast and efficient DMSO-mediated dehydration of carbohydrates into 5-hydroxymethylfurfural. Solenne D., Carole M,. Boris E. and al. s.l. : Elsevier, 2014, Catalysis Communications, Vol. 51, pp. 5-9.
19. Glucose conversion to 5-hydroxymethylfurfural with chromium exchanged bentonite and montmorillonite catalysts in different solvents. Aziz R.A., Solmaz A., Serkan N. K. 2019, Chemical Engineering Communication.
20. Increasing the revenue from lignocellulosic biomass: Maximizing feedstock utilization. David M. A., Sikander H. H., Shengfei Z. 5, s.l. : ScienceAdvances, 2017, Vol. 3. 28560350.
21. Toward biomass-derived renewable plastics: Production of 2,5-furandicarboxylic acid from fructose. Motagamwala et al. s.l. : Science Advances, 2018, Vol. 4.
22. Highly Efficient Two-Step Synthesis of 2,5-Furandicarboxylic Acid from Fructose without 5-Hydroxymethylfurfural (HMF) Separation: In Situ Oxidation of HMF in Alkaline Aqueous H2O/DMSO Mixed Solvent under Mild Conditions. Guangyu C., Linbo W., Hong F. and al. 48, s.l. : Ind. Eng. Chem. Res., 2018, Vol. 57. 16172-16181.
23. Towards improving the sustainability of bioplastics: Process modelling and life cycle assessment of two separation routes for 2,5-furandicarboxylic acid. Sara B., Pedro M. T., Eva R. 233, 2019, Separation and Purification Technology.
24. Solvent Effects on Fructose Dehydration to 5-Hydroxymethylfurfural in Biphasic Systems Saturated with Inorganic Salts. Yuriy R. L., James A. D. 52, 2009, Top catal, pp. 297–303.
25. A scalable carboxylation route to furan-2,5-dicarboxylic acid. Graham R. D., Amy D. F., Aanindeeta B. and al. 13, s.l. : Green Chemistry, 2017.
26. Enterprise Ionic Liquids Database (ILUAM) for Use in Aspen ONE Programs Suite with COSMO-Based Property Methods. Ferro V. R., Moya C., Moreno D. 57, s.l. : American Chemical Society, 2018, Ind. Eng. Chem. Res., Vol. 3, pp. 980-989.
27. PROCESS DESIGN PROJECT 2019 Manufacturing of Acrylic Acid. London : Imperial College London, 2019.
28. Catalytic conversion of corn stover for gamma-valerolactone production by two different solvent strategies: Techno-economic assessment. Byun J., Han J. s.l. : Elsevier, 2019, Vol. 175, pp. 546-553.
29. Sanan E., Paraskevi F., Agnieszka B. Direct Catalystic Conversion of Cellulose to 5-Hydroxymethylfurfural Using Ionic Liquids. Departement of Chemistry, Imperial College London. London : Andreas Taubert, 2016.
30. Triebl, Christoph. Simulation and Economic Analysis of 5-Hydroxymethylfurfural Conversion to 2,5-Furandicarboxylic Acid. Leoben University of Mining. Graz : s.n., 2012.
31. Estimation of Operating Costs. Chemical Engineering Projects. [Online] 11 May 2014. https://chemicalprojects.wordpress.com/2014/05/11/estimation-of-operating-costs/.
32. Prediction and screening of solubility of pharmaceuticals in single- and mixed-ionic liquids using COSMO-SAC model. Bong-Seop L., Shiang-Tai L. 7, s.l. : AIChE, 2017, Vol. 63.

Investigating Active Learning through Atrivity to Chemical Engineering

Katie Mawdsley and Komal Mehta

Department of Chemical Engineering, Imperial College London, U.K.

Abstract

Although the concept of active learning has been around for over two decades, it is only now being considered in the higher education of STEM subjects. This project aimed to provide the Department of Chemical Engineering at Imperial College with recommendations on the implementation of active learning using Atrivity, a gamified quiz app, into a first-year mathematics course. Preliminary research informed a series of trials, aiming to comprehend student preferences and learning styles, to design and test a question design framework which enhances knowledge acquisition, and short- and long-term retention, for use throughout the degree. Results show that students respond well to a progressive question design framework, perceiving this as helpful in their understanding and engagement. Students also enjoy the competitive gamification, but flexibility (through anonymity & opting out of time constraints) is essential. It was found that students derive the most value during tutorials with the option for further study in their own time. The implementation of active learning should be holistic, providing students with the resources to enhance their understanding of the syllabus, develop their core Chemical Engineering skills and bridge the gap between different pre-university mathematics qualifications.

Introduction

Active learning is facilitating the shift from traditional teaching of homogenously relaying information, to a truly interactive learner experience. Active learning aims to enhance the educational experience for both students and educators by using higher order thinking[1], harnessing the advantages of problem-based learning and gamified learning for retention[2]. It enables students to make connections among key concepts and facilitate the application of this knowledge in new settings[3]. It has been summarised to include four elements; talking and listening, writing, reading, and reflecting[4]. Although it can be argued that homework and coursework assignments offer active learning, it should be implemented within the classroom to be truly active[5].

The scope for active learning in STEM subjects is broad; typical activities and features include automated feedback and push notifications[6], regular clarification of notes with a partner[7], group worksheets[8] and gamified learning on mobile devices. Drawbacks include the associated costs in comparison to the cost effectiveness of traditional lectures[9]. Within higher education (HE), students are typically given freedom to manage their own learning and many universities now offer lecture recordings and other digital content[10], allowing students to engage differently with the material, increasing motivation and a sense of interaction[11]. It should be noted that the potential novelty factor which comes with new teaching methods, described as the Hawthorn effect[12], makes it important to test the effects in both the short- and long-term. In the literature, success has commonly been measured through exam performance, attendance[13], stress[14] and engagement.

This project follows the initial implementation of active learning into the first-year curriculum in the Chemical Engineering (CE) Department at Imperial College London. An app (by Atrivity - originally designed for onboarding activities in industry) was used during Chemistry lectures so that students could "challenge" each other in quizzes with Multiple Choice Questions (MCQs) written by their peers. The trial was pioneered by a new chemistry course lecturer using 'live games' during lectures. Active learning has also been implemented on similar platforms, such as edX[15] and Panopto[16]. Atrivity offers a unique opportunity whereby numerical data can easily be extracted for individual questions or users, which enables quantitative analysis, reducing the impact of bias which can often arise from interviews or focus groups.

There is a lot of debate in literature about the potential of MCQs and their ability to test from fundamental concepts to a much higher level of application; it has been suggested that they are still able to test a broad range of content appropriate for a full assessment[17], but other studies have highlighted the difficulty of composing MCQs which test high levels of knowledge and claim that sophisticated questions are better for this[18]. With regards to the material examined, MCQs allows for a broader range of questions to be asked rather than a small amount of questions which go into a lot of depth[19].

This research paper aims to answer the following:

How can an active learning quiz app be integrated into a first-year mathematics CE course to best prepare students for application of the concepts throughout the rest of their degree?

The study will look at the question design, game design and implementation of the app for a first-year Mathematics Fundamentals (MF) course to identify the optimum arrangement for both the students and educators. This project aims to provide recommendations in the implementation of active learning, considering both the short- and long-term effects; a gap in the literature currently. The department have identified active learning as an opportunity to engage students in the acquisition and retention of knowledge through a more informal environment where the consequences of failure are negligible. This prepares students for the use of concepts and tools throughout their degree, rather than solely the examination. The results include an effective question design framework and guidance on the implementation of the app into the course, through trials in lectures, tutorials and in students' own time. Data has been collected through numerous questionnaires, interviews, focus groups and extensive observations.

As students progress through their degree they are given the opportunity to specialise in different modules, with individual pre-requisites, but the theory and application taught in the MF module is uniquely applicable for, and underpins, all other higher level modules unlike the other fundamental courses, motivating the study of this course. It is designed to result in a 'level playing field' such that all students have a similar ability upon completion. The

initial differences arise from variations in pre-university maths qualifications, extended further if a student has not previously studied a 'Further Maths' (FM) (or equivalent) module, resulting in a significant range of prior knowledge. This differs from other fundamental courses, where most concepts are introduced for the very first time. The Institution of Chemical Engineers (IChemE) accredits degree courses reaching and maintaining high standards; they set learning outcomes and define the capabilities individuals must obtain to be awarded with their degree. One key learning outcome is the underpinning mathematics, science and associated engineering[20]; thus, highlighting the importance of the mathematical ability of a Chemical Engineer. IChemE degree accreditation requires the understanding of mathematics necessary to support the application of CE principles and processes. The MF course taught to first-year chemical engineers is designed to provide students with tools and concepts vital for progression through the course. It has 55 lectures across two terms, alongside 18 tutorials and examined via one three-hour examination at the end of the year (see A.2 for the syllabus). Rosenthal[21] trialled various active learning methods in HE maths lectures and found that they lead to deeper understanding and better retention of knowledge by the students.

Background

Petty[22] described how the process of actively learning through "making or doing things" allows learning to be more efficient. Other research has suggested that our brains require physical activity to learn[23], cognitive processes are best learned when robustly practised[24] and "expertise results from deliberate practice"[25]; all enabled by active learning. Joosten[26] found that active learning can both increase engagement and retention, and more specifically, it has been shown that Engineering students will remember more content if brief activities are introduced to the lecture[5].

Although there has been limited research specifically applying to CE in HE, The Study of Binqui[14] found that using gamification on a smartphone app, incorporating the problems without data methodology, reduced stress levels and increased performance in a particular CE module. Studies have also found that when faced with a problem, CE students search for a formula to guide them towards the solution[27], highlighting a need for a bridge between acquisition and application of concepts in alternative settings. One study on a CE module found that the implementation of active learning through seminars, prepared and delivered by students, and case study activities saw good student engagement with the subject material[28], but there is limited knowledge on the longer term effects and whether students are able to apply their understanding to unfamiliar problems. Freeman et al.[29] also found that undergraduate performance in STEM courses increased by 0.47 SDs under active learning, through meta-analysis.

A correlation between active learning and achievement is yet to be widely established but research suggests that it incentivises collaboration and social engagement[31, 32]. The vague measure of 'engagement' is used extensively in pedagogical research with frameworks featuring behavioural, sociocultural, psychological and holistic measures[32]. The rewarding aspect of gamified active learning drives internal curiosity and reflection[32], showing the benefits of increased engagement. Furthermore, gamification can "reframe" failure, important in the learning process[34].

The design, delivery and implementation of educational materials are usually at the educator's discretion and is often motivated by personal preferences. One learning framework from Meyers and Jones[4] insists on a structured and organised implementation of active learning in the classroom where a student would "clarify, question, consolidate, and appropriate new knowledge". A significant limitation of this framework comes from the structured approach limiting potential creativity. Wilkinson[32] proposed a framework based on Mobigames for anatomical sciences in HE for the integration of active learning. The framework features four distinct aspects: information, facilitation, learning and timing. This framework encourages the educator to consider the whole learning cycle, highlighting the importance of educator facilitation and the consideration of the timing of implementation. A third framework is Bloom's Taxonomy[35]; represented as a pyramid with the base being 'remember', working upwards to 'understand', 'apply', 'analyse', 'evaluate' and 'create' (see Figure 2). The levels have been mapped to cognitive skills by Anderson & Krathwohl[36], and simplifications of the number of levels have also been deemed an appropriate assessment of understanding[17, 37]. Frameworks incorporating various knowledge levels are very common and used extensively. A simplification of the knowledge levels are 'recall', 'application' and 'evaluation'[38], demonstrating its flexibility. Similarly, the IChemE defines 'areas of learning' as 'awareness', 'knowledge', 'understanding', 'skills' and 'complex' (implying engineering problems and processes). Within the realm of Engineering education, one widely used framework is the "Conceive, Design, Implement and Operate" (CDIO) framework which stresses the fundamentals and bridges the gap between Engineering education and real-world demands[3], explicitly featuring active learning. Aspects of these models will be used to inform the question design and implementation of the app in the MF course to ensure it delivers a high quality of learning potential for students. These frameworks will also assist in defining metrics for the measurement of difficulty of questions. The level of question will be correlated with student performance in each question to be able to distinguish, and ultimately allow students to assess their understanding of topics by their performance on each level of question.

Methodology

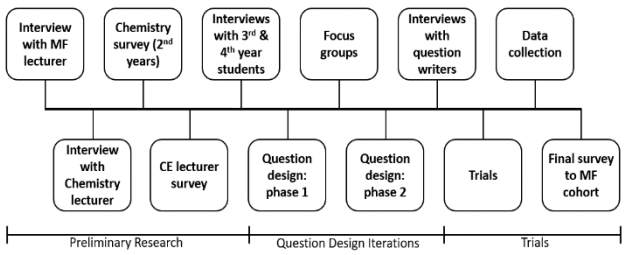


Figure 1: Methodology timeline

The main research question was split into four categories:

- Question design
- Game design
- Implementation
- Effects of pre-university maths qualifications

All interviews and questionnaires were conducted following the appropriate ethical guidelines, providing privacy notices (including anonymity disclaimers) on surveys, and consent and information sheets for focus group and interview participants.

The MF course leader was interviewed to understand the syllabus, delivery and examination of the course. A secondary objective aimed to comprehend their views on active learning for the MF course. This aimed to identify any potential opportunities to enhance the educator's or learner's experience. All interview questions (A.1) were open ended and neutral, singular and clear[39] to avoid confusion and to collect the desired information without omitting the discussion of topics which had not been previously considered.

Another interview was conducted with the course leader who led the trial in the Chemistry module with the objective of understanding the logistics behind the trial, question design, how it was implemented, and any general results and observations. The lecturer was also asked about which features of the implementation worked well and not so well, to identify opportunities for further testing, or inform this trial to avoid some features. The interview was designed in accordance to Rosenthal's recommendations discussed above (A.3).

The next stage involved recognising the needs of the students by conducting a questionnaire (A.4) for now second-year students to review and reflect on their experience using the app during their first-year Chemistry course. An online questionnaire was used because this was most accessible, so that a significant number of responses could be collected in a short amount of time. The questionnaire was advertised solely to second year students and was sent to their mailing list and a social media page. Students were also encouraged to participate in person during their laboratory sessions. Questionnaire responses are easier to analyse and interpret because they fall into a limited set of categories[40]. Some questions were inspired by responses from the course leader, for example, asking whether students used the app and why, and their perception of the usefulness and the different features. To gauge the usability, students were asked when or where they used the app, with available responses being in, before or after lectures, during tutorials or during the revision period; students could select multiple responses. The 'other' option was also given along with a text box to elaborate on any options that had not been previously recognised. The next question extended this, where students were asked to select one of the above options to gauge where students saw the opportunity to derive the *most* value from the app and was manageable within the first-year schedule. After reviewing the previous feedback and preliminary brainstorming, a selection of additional features was identified which formed the basis of the response options for the features question. An 'other' option was offered again to identify new options. Students were then asked to indicate which courses they felt would benefit most from having the app, since they had already experienced using it. Following this, they were asked two questions about their personal learning preferences; whether they were motivated by the competitive nature of the app and whether they perform well under pressure. The response options were derived from a version of the Likert scale using 'strongly agree', 'agree', 'disagree' and 'strongly disagree'[41]. The omission of a neutral option prevents the 6-23% of people who are likely to choose it[42] and encourages people to report meaningful attitudes they would otherwise not be bothered to share[43].

In addition to the questionnaire, numerous face-to face interviews were conducted with CE students relating to their experience with the MF course. The students were asked open questions to gain insight into the link between experience and performance (see A.5 for questions).

A questionnaire was created in the same style as the second-year questionnaire to collect the views of course leaders for all courses beyond the first year (A.6). This asked them to indicate which MF topics are pre-requisites for their module and asked for additional comments about the mathematical ability of students. This information was to be paired with a previous questionnaire completed prior to this project, which asked students to express the level of learning of the MF topics they had previously covered prior to University[48], to identify critical topics, providing the basis for gauging whether students were sufficiently prepared for both their end of year exam and the rest of their degree.

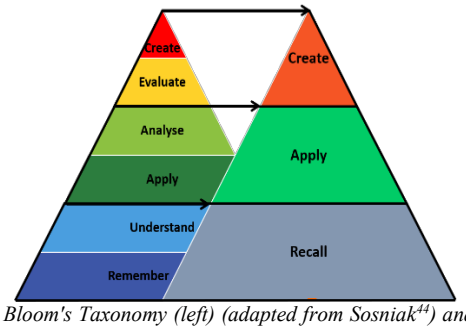


Figure 2: Bloom's Taxonomy (left) (adapted from Sosniak[44]) and the modified version (right) forming the basis of the Question Design Framework

Having gathered the preliminary data, the next step was designing the questions. A modified version of Bloom's Taxonomy was selected: 'Recall', 'Apply' and 'Create' (Figure 2). Using a three-tiered method, rather than six meant that students would more easily be able to distinguish between the levels. It would also make the question design a little simpler, using the framework as a guide which still offers flexibility. This framework coincides with the IChemE's learning guidelines omitting the first and last (awareness and complex), since awareness comes from attending the course lectures, and complex arises further into the degree, rather than in parallel with the learning. 'Recall' questions are based on the theory within the lecture notes and 'Apply' tests students on the use and application of the knowledge in different contexts. 'Create' simply asks students to create a question based on a MF topic for use within the app.

Initially, twelve questions for two topics which had already been taught were designed, acknowledging that this part of the methodology is an iterative process, requiring input and feedback. There were six recall and

six apply questions to be tested within a focus group, each with three MC responses.

The app can release games on a predetermined schedule and the number of questions for each game can also be defined which was set to 7 for the trials. The two main game types are 'challenge' and 'training'; both were enabled throughout the trials. The first allows players to choose an opponent to play against or select 'auto-match' where the opponent is randomly selected. 'Training' allows the player to answer questions without competing against another player. The time per question was always set to 1000s and points were rewarded for correct answers only and not the speed in which they were answered.

The focus group was designed to test the question and game design, in accordance to Kreuger's[45] progressive focus group design, using the final third of the focus group to engage participants[46] (A.7). Participants were recruited by sending an email to the entire year group asking for volunteers, with a voucher incentive and refreshments throughout the session. Prior to introducing the app, a discussion about the MF course encouraged the students to reflect on their experiences versus their expectations. The students were then asked to play a game on the app and take note of the difficulty (on a scale of 1 – 5), flexibility (can be done without paper), advancement (reliance on notes) and how long it took to answer each question. The students then answered questions about their thoughts on the game design and opportunities and drawbacks associated with the app. The 'create' element of the framework is aimed to stretch students further and this was initially tested in the focus group by asking the students to pick one of the two topics they had covered in the MF course and write a question suitable to be added to the app; the students were asked to share their thoughts on the usefulness of this process. Beyond this, students who had been recruited during the Chemistry trial to design questions were interviewed to understand their experiences (A.8). The students were asked if they followed a framework, talked through the typical process they followed when writing questions, discussed what worked well and not so well, and identified the easier and more difficult parts of the process. Each question was open ended to facilitate reflection and discussion to understand their experiences and the potential of this method being rolled out to all students. It should be noted that these students had been paid to write questions.

Following the focus group, the question levels were redefined to ensure each level reflected the correct knowledge and skills, and several changes were made to the game design. A second focus group, designed in the same way, was then held to make further refinements.

Before starting the trials, the first-year cohort were requested to download the app by the course leader (emphasising its importance) on the virtual learning page, during the lecture and by email. This was to ensure each student received the message at least once. Other reminders and announcements were sent out from the departmental expert of the app. Four trials were completed; twice in tutorial classes, once in the students' own time during a reading week and once during a lecture. Games on topics which had already been taught were available to play, meaning that students could test their ability and identify gaps in their learning, highlighted as an important step during the focus groups. During the tutorials, the tutors were given instructions to assist students to login and choose which topic to play, leaving students for up to ten minutes whilst observing their engagement. A questionnaire was sent to the tutors upon completion of both trials to collect their thoughts (A.9). During a reading week, a task was set via email to play two games before the end of the week. The final trial was held in the lecture theatre, where the lecturer asked students to play games with each other on topics which had already been taught. In each setting, the game set-up was the same: seven questions and a 1000 second time constraint per question to eliminate the time pressure, allowing students to compare the experiences.

Data was extracted from the Atrivity control panel to analyse the overall uptake of students for challenges and trainings for each topic. The data provided by the app includes when students used the app, the number of times specific questions were answered incorrectly, and the type of game they played (training or challenge). Further in-depth data is provided by Atrivity such as 'Knowledge Level' and '% Engagement', however this was not used as the formulas had not yet been provided.

Upon completion of the trials, a final survey (A.10) was conducted to gauge the student opinion on the app. To maximise the number of responses, tutors were asked to allow five minutes at the start of the tutorial for students to complete it, rather than relying on them to do it in their own time. Graded response options 'strongly agree', 'agree', 'disagree' and 'strongly disagree' were used[41] including a neutral option to account for students who had not used the app in a particular setting, or did not believe the app had made a difference to their experience. The final question asked students to indicate their level of pre-university maths qualifications.

Results & Discussion

Preliminary Research

The MF module leader, while being interviewed, admitted reluctancy in implementing active learning through an app due to believing it would "prevent students from handwriting notes and trying the problem sheets, which are key to the learning process". However, after observing reduced engagement from students, partnered with low lecture attendance due to either students having seen the material before and hence it being too easy, or others complaining that the course material is taught too fast, they felt ready to welcome a "modern" element to the course. This is consistent with the results from the first-year survey (A.10), indicating 18.8% of students consistently do not attend lectures. The lecturer identified two main benefits expected: offering students the "chance to take risks whilst the stakes are low" and introducing an "element of fun into the module", which reduces the reliance of students on bulimic learning[47], enabling students to retain core concepts for longer term application. The lecturer also identified an additional opportunity in "providing students with an extra resource to bridge any gaps" between differences in pre-university qualifications, motivating further exploration within this research.

The lecturer who pioneered the trial in the Chemistry course conveyed, during their interview, that due to the constraints of implementing the app whilst designing and delivering new course material in parallel, the trial did not engage a representative number of students to draw meaningful conclusions. Feedback suggested that students derived the most value from conceptual questions and the lecturer found "the use of the app in lectures helpful in identifying gaps" in learning and understanding, however students complained about the repetitiveness of questions. Students were recruited to write questions in parallel with the course once the material had been covered in lectures; the students did not follow a framework but were briefed to write MCQs which tested conceptual understanding. Each question was designed to have one correct answer and four incorrect answers which could be rotated to be used multiple times, which has since been identified as the cause for complaints, because although the possible answers changed, the same question was being asked. The lecturer admitted that "it was difficult to engage students to write questions" and recommended asking students who are more comfortable with the content (i.e. higher year students).

The app was offered during lectures and during the revision period for students to use in their own time; the latter saw a higher uptake, highlighting the potential of the app as a revision resource. At the beginning of the course, 36% of students were active in the app (in line with low lecture attendance), which dropped to 13% by the end of the course[48], however this is believed to be due to external factors (such as disengagement with material due to complexity) and not a direct result of the app. The students enjoyed the competitive aspect but not the time pressure and said that "the app offered limited learning potential due to the lack of solutions". The final grade for the students who played the game during the revision period was 73.9% (SD 13.8%) and the average mark of the remaining students who did not use the app was 66.2% (SD 21.9%)[48]. However, it is possible that the more proactive students used the app and are more likely to use all resources available to them and so will be high achievers regardless of the availability and usefulness of the app, motivating the need for further research.

Out of the older year students who were interviewed, 75% said they would have used the app for MF to consolidate material and those who said they wouldn't said it was due to time pressure in the first-year curriculum. Contradictory to this, one student said, "it'd be convenient considering how busy we were, as it is accessible and would stop me procrastinating on my phone", consistent with findings in the literature[49].

Question Design

Question levels 'Recall' and 'Apply' were tested by recording the perceived difficulty and time taken to answer from focus groups, and failure rate taken from automatic analysis completed through the app. One important preference identified through the focus group was students preferring the difficulty of practise questions to build up as they progress, suggesting that a question design framework which had progression levels

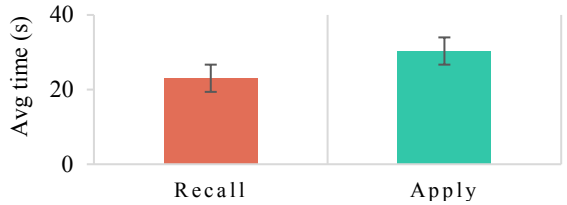


Figure 3: Average time taken (s) for students to answer questions (Focus Group)

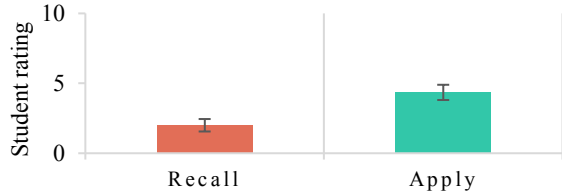


Figure 4: Average perceived difficulty for questions (Focus Group)

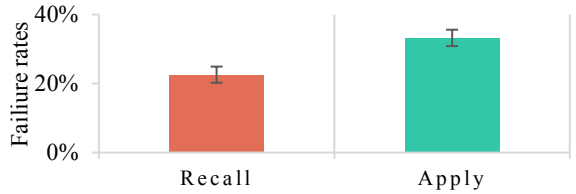


Figure 5: Average failure rate (%) for questions (Atrivity Analysis)

would engage students the best and so the tests sought to understand whether performance reflected the levels.

The first test looked at how long students took to answer the questions; the average time taken to answer recall questions was 23s and apply questions was 30s (Figure 3). Overall, any time taken under one minute was deemed appropriate in this study for students to keep track of their thoughts without notes or losing focus. The additional time taken for students to answer application questions was as expected because the proposed framework is hierarchical, thus, to solve an application style question, students must also demonstrate the mastery of the levels below it, i.e. students must recall the knowledge before applying it to the question[50]. The perceived difficulty of application questions (4.4 out of 10, where 10 is very difficult and 1 is very easy) was also higher than recall questions (2 out of 10), again as expected (Figure 4). Both difficulty scores were relatively low, suggesting that the questions were not very challenging, but it is important to note that the questions used during these tests were on topics which typically students find less difficult, and so once more difficult topics have been covered in lectures, it is expected for these averages to increase. Students also expressed a much greater need for their notes when tackling application questions.

Data was retrieved from the Atrivity app which measures the number of times an incorrect answer is submitted for each question, defined as the failure rate. On average 23% of answers to recall questions and 33% of application questions were submitted incorrectly (Figure 5), suggesting students struggle with the application questions, as less students arrive at the correct answer. The highest failure rate recorded within the trial was 94%, an application question; this highlights the need for educator (lecturer or tutor) engagement where they can check the data and provide extra support on specific topics or methods.

On a more qualitative basis, students were asked in the final survey whether they believed the app had the

potential to: increase their engagement, consolidate their learning and increase their confidence (Figure 6). In all three cases, most students agreed or strongly agreed.

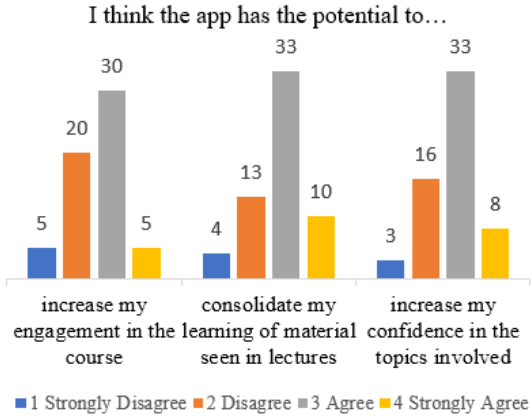


Figure 6: Number of students who strongly agree, agree, disagree or strongly disagree with each of the statements above (Final Survey)

There were a significant number of students who disagreed or strongly disagreed, which may be down to differences in student learning styles or implementation between tutors, whether the student participated in each trial type (own time, tutorial and lecture) and their pre-university maths qualifications, meaning that they may have covered the material previously and so using an app is not perceived helpful. The focus group students who had not studied FM all said that they would like to play games which covered content bridging the gap between their studies and what is assumed knowledge during the course. Due to the short time frame, questions relevant to the whole cohort had to be prioritised in the interest of collecting meaningful feedback so these questions were not used within trials but will form part of the recommendations. Focus group participants were unsure whether the app had the potential of helping them with the "really hard" questions, but this may be due to having only been authentically exposed to the first two levels of questions from the proposed framework. The app does have limitations in its functionality, and, if the app is to remain flexible so that it can be used, for example, on a commute, then the questions will have to stay relatively simpler than examination questions. Students also complained that they had been unable to re-attempt a question they had incorrectly answered originally, in line with observations seen in literature, which recognises importance of attempting a question again[51].

The third and final level of the framework, 'Create', was originally tested in the focus group by asking students to prepare questions and three MC solutions. This was designed to stretch and encourage further engagement with the course material. During the focus group, all students expressed that they did not yet feel comfortable enough with the material to write questions, but also said that they had little confidence in this being helpful, explaining that they would not be motivated to write questions and would skip it. Following the focus group, students who had written questions during the chemistry trial were interviewed and said, "writing questions did not help retain knowledge through to the examination period". The student-writers did, however, find that "writing solutions, and especially identifying how students may take an incorrect step when solving a problem, was helpful". This assisted the students in remembering subtle points within the material, and better prepared them for their exam. Asking students to write solutions to questions in their own words to be submitted to a solution "bank" was considered, but the practicalities in offering unique feedback to each individual meant that this would introduce a lot of new work to the educator, which may be infeasible. This motivated the use of a new type of question, in which students are asked to identify incorrect steps in a proposed solution and identify the method to correct them. These questions can be set in advance and be repeated much easier, preparing the students for their examination in being more aware of the typical mistakes. These questions also have the potential to be inspired from previous examination responses, which the lecturer can flag as common mistakes and can be integrated into the app, offering the students a type of pre-empted feedback in the form of suggestions of where they may go wrong. An example of a question is provided in the appendix (B.2) but has not been fully tested yet.

In addition to the three-tiered framework proposed and tested above, lecturers expressed a concern for students' lack of comfort in performing estimations and order of magnitude analysis in the lecturer survey, something which is required for many lecture-based and project-based modules. Although not tested thoroughly, some questions were designed (B.2), and will form a part of the recommendations, to offer students the opportunity to develop their wider CE toolbox beyond the MF syllabus. Gianni Astarita[52] states that "the techniques of dimensional analysis, appropriate scaling and intelligent estimation of orders of magnitude constitute an extremely powerful tool" for all CE and most information about a problem can be analysed with this tool.

Game Design

Preliminary research found that 68% of students find the competitive element of the app motivating (C.4). Therefore, the 'challenge' aspect of the app is ideal for most students. However, it is important to cater for the 32% of students who are not motivated by the competitive element, so a 'training' option was enabled on the app for students who don't want to compete against other students. However, 324 challenges were played compared to 67 trainings during the trial,

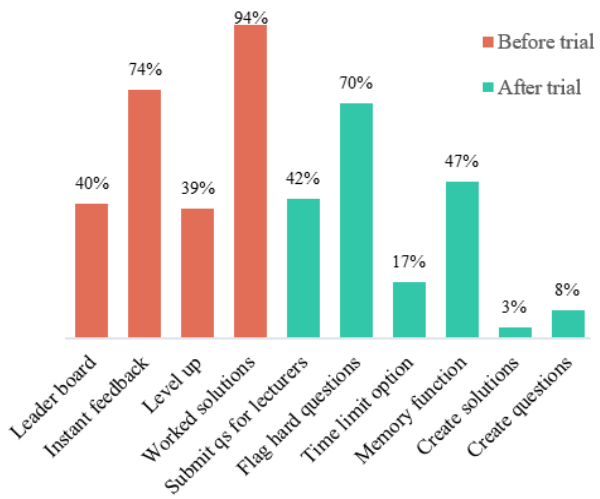


Figure 7: % of students that prefer each of the features listed, before (2nd year student survey) and after (Final survery) the trial

demonstrating how most students do enjoy the gamified aspects of the app (C.5). A key feature which 94% students selected as desirable was access to worked solutions (Figure 7); a key aspect of student achievement[53]. This feature has not been integrated due to constraints in the app but use of solutions were tested during the focus groups. Prior to using the app, 75% of participants in the focus group expressed that they had resorted to external resources, although sometimes overwhelming, to enable their understanding and seek worked solutions; the same students agreed that additional material in the form of questions and solutions, especially endorsed by the course leader would be preferred. Upon use of the app, 75% of students expressed that they were able to identify their strengths and weaknesses, whereas the other 25% of students had not been able to differentiate which questions were associated with which section of the syllabus. Due to the limitations of the app, section references were not able to be added but will form part of the recommendations following this report. It is important to note that many students will be tempted to retrieve solutions and bypass the use of the app; in order to avoid this, the solutions must be integrated into the app so that they are only available after every game, and students shouldn't have unlimited access to them. Instant feedback was selected by 74% of 2nd year students and is a feature that was readily enabled in the app by indicating correct and incorrect answers to each question (Figure 7). During the focus group, students suggested having an option to be anonymous on the leader board so that they could use this as a metric to measure their progress but without others knowing their position and so this was implemented, such that students could opt in or out of doing so, and also catered for the students who did not find the competitive element of the app motivating. During the trial, 30% of active students changed their name to be anonymous on the leader board (found from final survey), representing the 32% of students who are not motivated by the competitive nature of the app (C.4). Careful consideration should be taken when students are using the leader board as a metric of performance and progression, because scores are not correlated to ability. Students should be coached on the false sense of security, or opposite, worry that the app may cause and so should assess holistically, such as through performance in progress tests and ability to complete problem sheets.

After the trials, students were asked the same question regarding desired features, inspired by ideas generated in the focus group. The most prevalent feature was flagging difficult questions, for which, once a question is flagged a certain number of times the lecturer or tutor can address it with the class. 42% of students selected 'submitting questions for lecturers', similar to flagging, highlighting the importance of the involvement of the educator (albeit lecturer or tutor). Another popular feature is a memory function which records the questions that have already been answered, at 47%, so that students can request to study only the questions they have previously answered wrong, again, showing the importance of personalisation. Regarding the time constraints set in the app, focus group participants explained that they "prefer being able to work at their own pace when learning a topic, but nearer to the exam, time constraints may become more helpful" as they begin to increase their pace to be able to complete the exam within the given time. Although this study will not extend to exam preparation, this highlights that the design and implementation of the app may change significantly depending on when the app is used. This applies to question design and time constraints.

Implementation

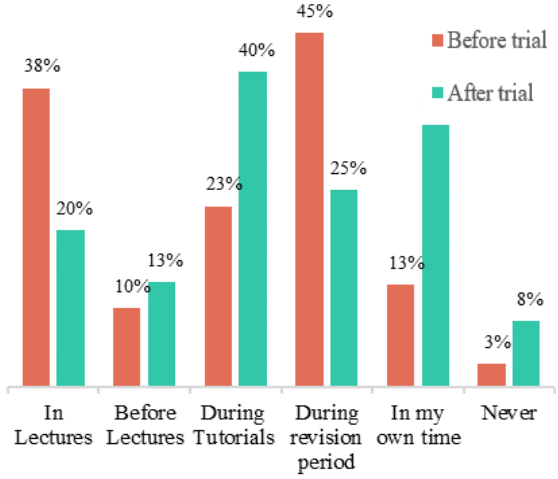


Figure 8: % of students that preferred each of the settings above, before (2nd year survey) and after (final survey) the trial

During last year's chemistry trials, 40% of the students who used the app at all, used it during their revision period and all used it during lectures as it was compulsory (2nd year survey). When asked their preference of usage, 45% would have preferred to use the app during revision periods, 38% during lectures and 23% during tutorials (Figure 8). Additionally, although the focus group students expressed their appreciation for tutorials, highlighting this as their main port of call, they agreed that tutorials offer limited added value. These comments motivated the use of tutorials during the trials, as an opportunity to enrich the learner experience by equipping tutors with an interactive app, utilising unseen content, whilst using the other time to focus on questions students particularly need help with. Additionally, one student admitted that they "only copy the answers down in tutorials", having not previously attempted the problems and is unable to follow, and so expressed an interest in using the app to attempt questions at a level (or topic) they were more comfortable with, which has the potential of personalising the tutorial experience for each individual student, without compromising others or the tutor. As a result, it was decided that each trial would incorporate a choice of three games, each with a different topic. Due to the question design being an iterative process, the question levels had not yet been tested rigorously enough to split by level, and in addition, the number of questions written per level would not yet form a rich enough selection for students to derive value. After studying the uptake of each topic, in both batch one and batch two, the most popular games were with the topics which had been covered longest ago (C.5), rather than the recent topics, highlighting students' preferences to consolidate learning before attempting questions.

As the three most popular settings were in lectures, in tutorials and during the revision period, these would have been the settings used, but the timeline of the trials meant that there would not be a revision period and so this was shifted to be during the student's own time during a reading week. Following the trials, students rated each of the settings for which the results are shown in Figure 9.

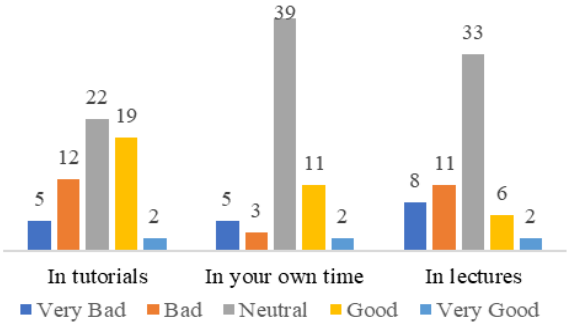


Figure 9: Number of students that gave each of the ratings for the trial settings, from Final Survey

The graph shows a very skewered response, reflecting very different learning preferences. It was consciously decided to offer the neutral option for the students who did not attend lectures or tutorials and due to the very low uptake of students playing in their own time. The wording chosen, upon reflection would have been better as "N/A" to avoid students being dishonest and picking neutral for ease.

Students were also asked which setting they would prefer to use the app in, in the future, yielding more conclusive results. Tutorials was most popular, with 40% of students selecting it as their preferred setting (Figure 8), coinciding with tutorials having the highest number of student ratings of 'Good' or 'Very Good'. Not all tutors engaged with the trial, which may have been due to the lack of training, or them not believing in its potential, and so if the app was implemented with more consistency and enthusiasm, a higher rating in tutorials would be expected. Observations saw disengaged tutors spending the time whilst students were on the app on their phones, rather than engaging with students, which also lead to students switching focus and using their phones for social media usage. Comments from most of the tutors through the survey supported the use of the app in tutorials and highlighted that students enjoyed the competitive aspect of the game; "Students were engaged and were explaining/helping each other when they had problems with a question" and "I think they especially enjoyed challenging their friends in the tutorial". The implementation in tutorials would require a detailed plan of a schedule, as one tutor suggested "If the app were to be included in tutorials I think it would have to complement the tutorial questions and tutorial/app should be used together - or maybe as a challenge at the end of the tutorial to test their knowledge".

Although 49.3% of students did not use the app in their own time during the trials (C.3), after the trial, 33% of students said they would prefer to use the app in their own time (Figure 8). The low uptake could be due to external factors as the first-year students have a busy schedule. The decrease in students preferring to use the app during the revision period could be due to the question and game design not being geared towards exam revision but focusing on core concepts. Upon completion of the trial, students were less interested in using the app during lectures, which may be due to time constraints or lack of engagement with the educator during the trial. The lecturer also expressed that there is not space in the current curriculum for the app to be used in lectures.

Pre-university Mathematics Qualifications

Prior to using the app, when asking how the focus group participants were finding the course, it immediately became apparent that the experience is dependent on whether or not the students had studied FM (or an equivalent) previously; with each individual student beginning their response with the statement of whether they had studied it, before going on to make other points. From the interviews with the older year students, not a single student who had not studied FM achieved a MF grade higher than their average for that year, whereas 70% of students who had studied FM had done. This immediately highlighted an opportunity to offer a resource to bridge this gap.

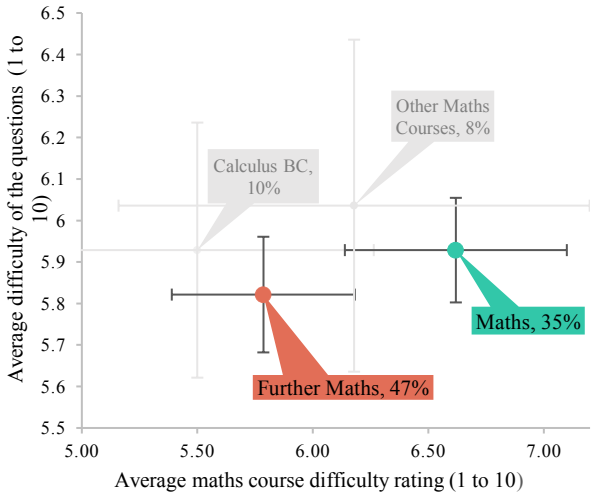


Figure 10: Average MF course difficulty rating and average app question difficulty rating for different pre-university maths qualifications

Figure 10 shows the differences in perceived difficulty of both MF and the questions in the app between students who studied FM and those who didn't. Students who didn't study FM find the maths course harder than those who did, as expected. However, the perceived difficulty of the app questions is almost identical, highlighting the accessibility of the app for all students. Only 17% of students had qualifications outside of the A-level & Equivalent category and had significant error, so were omitted from the analysis. Finally, from the second-year survey, it was found that only 67% of students who had studied FM would have used the app if it had been for MF, whereas 80% of the non-FM students said they would have used it. This data provides a better understanding of pre-university maths qualifications, however student performance for the current cohort is hard to measure as there were no exams during this period. The differences found provides a good incentive for the active learning approach through which the knowledge gap can be bridged in an efficient way.

To prepare and equip students with the relevant knowledge and skills, input from module leaders regarding their mathematical prerequisites was collected (A.6). A total of 30 survey responses were collected from

lecturers selecting the required topics to understand the material within their course. Responses were then combined with a previously conducted survey[48] sent to first-year students to gauge the extent to which students had previously covered these topics prior to university. The percentage of courses requiring each topic, and the percentage of students who had never covered the topic are plotted on Figure 3 below. The colour is a measure of the critical nature of the topic, demonstrating that most mathematics fundamentals topics require teaching to begin from the basics and progress to a level of complexity which will prepare students for application elsewhere in the course, highlighting the criticality of both knowledge acquisition *and* retention. This also highlights the importance of bridging the gap between differences in pre-university qualifications.

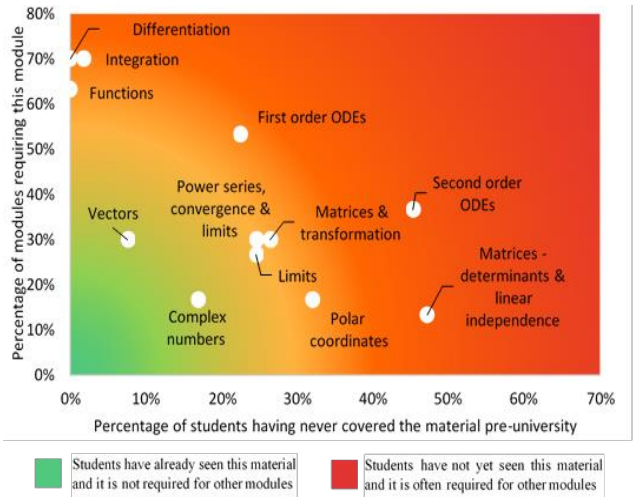


Figure 3:For the MF topics, the % of students who have never covered the material pre-university and % of modules that require the topic.

Conclusions

Overall, students trust and respond well to the question design framework which analogous to the IChemE framework. The performance of students during the focus group and in the app complements the different levels as expected and 72% of students are confident that the questions consolidate their learning. The focus on conceptual questions as well as CE questions is useful for students to develop their wider CE toolbox.

The numerous limitations with the features and flexibility within Atrivity is a downfall, all of which can be resolved in the long-term. However, inbuilt features that students responded well to were the leader board and challenges. Students engage with the competitive nature of the app; one of the tutors observed that there was "Lots of fun and joy when students competed against each other". This is supported by the 60% of students who find the competitive element of the app motivating (C.4). All students must be catered for, which is why flexibility within the app to opt in and out of time constraints should also be available. A crucial feature of the app, yet to be integrated, is worked solutions, with 94% of students selecting this as a desired feature.

The most effective way that the app can be integrated into the MF course is through tutorials as well as allowing for students to access games in their own time or during revision periods, with 73% of students selecting at least one of these settings as their preferred choice. It is important to note that uptake was considerably higher for topics covered earlier in the year than more recently, determining that students are using the app as an extra resource to consolidate material rather than acquire.

Differences in pre-university maths qualifications does affect a student's experience with the MF course in their first year, highlighted by the results from interviews with upper year students which show 0% of those who didn't study FM did not achieve above their average grade in the MF course. Students who haven't studied FM are also 13% more likely to seek other resources to assist in their learning, such as the app.

Recommendations

Following the results and discussion based on this project, the following recommendations are proposed:

- Questions should be designed following the three-tier framework: 'Recall', 'Apply' and 'Create'.
- The app should offer students the following features; option to submit or flag questions and immediate, but not unlimited, access to worked solutions after each game. It should also allow flexibility with challenging other students, time constraints and the appearance on the leader board.
- Active learning via use of the app should be implemented into tutorials, modifying the structure of tutorials so that the focus is not solely on problem sheets. The games played in tutorials should be related to the same topic as the problem sheet.
- Topics should be released for students to play in their own time after tutorial sessions with harder questions becoming available on completion of the initial batch. Release times can be set well in advance on the Atrivity website. A range of games focusing on building the student's Chemical Engineering 'toolbox' should be available for students to access, involving estimation, dimensional analysis and order of magnitude practise.
- Games should be available for all students but some select games should be tailored for students who did not study Further Maths, which focus on the knowledge gaps that exist and provide an extra resource for students to practice in their own time.

Limitations

There are many limitations which are associated with pedagogical research which should be monitored to assess their impact and minimise effects.

The first is the importance, and therefore reliance on the self-evaluation of students, but with ranging learning styles, preferences and expectations, the same response from two different students can be based on very different metrics, leading to a lack of reliability[54]. Additionally, this study worked closely with first-year students who do not have much experience in HE to compare to, resulting in their responses and suggestions sometimes being impractical. The subjects of pedagogical research are humans, and so control variables cannot always be controlled, requiring additional observations and

extensive interviews to understand the wider context. It is also unethical to prevent access to a group of students to act as a control group, making it difficult to draw conclusions without making assumptions. Repeat 'experiments' are also extremely difficult to conduct, due to students constantly progressing and moving on, or waiting for the new academic year to begin. For this project, the next steps should be to repeat the tests on the different question levels zero, three and six months after the original learning of a topic to test if student performance consistently reflects the difficulty levels and whether knowledge is retained. There are also significant limitations in using the Atrivity app, with many recommendations being unimplementable with the current setup; some of which have been discussed within the report. A major limitation of Atrivity is its incompatibility with equations and so they cannot be inputted as text, and must instead be replaced with an image, but this can often be of low quality or so small that the images are hard to read. Students also commented on these images making the app look underdeveloped and of low quality compared to other apps used on their smartphones (C.1 /C.2). There are also strict character limits on the questions (200) and answers (100), making it difficult to summarise complex questions and provide context for wider application questions. Hints and solutions are also unable to be accessed in situ, with students having to be directed to alternative platforms to access these at the end of a game, disrupting their app experience.

Acknowledgements

We would like to thank Dr. James Campbell for all his direction during this project. We also would like to thank Dr. Deesha Chadha for her guidance on both ethics and pedagogical research and Prof. Serafim Kalliadasis for supporting the trials. Finally thank you to the lecturers, students and tutors in the CE department for their input.

References

(1) Bonwell CC, Eison JA. Active Learning: Creating Excitement in the Classroom. *ASHE-ERIC Higher Education Reports.* 1991.
(2) Pechenkina E, Laurence D, Oates G, Eldridge D, Hunter D. Using a gamified mobile app to increase student engagement, retention and academic achievement. *International Journal of Educational Technology in Higher Education.* 2017.
(3) CDIO. *CDIO Standard 2.1.* Available from: http://www.cdio.org/content/cdio-standard-21 [Accessed 02/12/2019].
(4) Meyers C, Jones TB. *Promoting Active Learning. Strategies for the College Classroom.* Jossey-Bass Inc, San Francisco; 1993.
(5) Prince M. Does Active Learning Work? A Review of the Research. *Journal of Engineering Education.* 2013.
(6) Garbrick AH. *Factors influencing student engagement in an online asynchronous discussion forum measured by quantity, quality, survey, and social network analysis.* Doctor of Philosophy. The Pennsylvania State University; 2018.
(7) Ruhl KL, Hughes CA, Schloss PJ. Using the Pause Procedure to Enhance Lecture Recall. *The Journal of the Teacher Education Division of the Council for Exceptional Children.* 1987.
(8) Scherr RE, Hammer D. Student Behavior and Epistemological Framing: Examples from Collaborative Active-Learning Activities in Physics. *Cognition and Instruction.* 2009.
(9) Christersson C, Staaf P. *Promoting active learning in universities.* European University Association. Report number: 5, 2019.
(10) Hiltz SR, Turoff M. The Evolution of Online Learning and the Revolution in Higher Education. *Communications of the ACM.* 2005.
(11) Fernandez V, Simo P, Sallan JM. Podcasting: A new technological tool to facilitate good practice in higher education. *Computers & Education.*2009;53(2):
(12) Fraij F, Al-Dmour A. Effect of visual-assisted pedagogy for teaching and learning properties of relations. *International Journal of Technology Enhanced Learning.* 2014.
(13) Brown A. Using evidence-based active learning approaches in Engineering: How does it impact engagement and performance? *2018.*: STEM Conference/
(14) Paniagua S, Herrero R, Isabel Garcia-Perez A, Fernando Calvo L. Study of Binqui. An application for smartphones based on the problems without data methodology to reduce stress levels and improve academic performance of chemical engineering students. *Education for Chemical Engineers.* 2019.
(15) edX. *Engineering Courses.* Available from: https://www.edx.org/course/subject/engineering [Accessed 08/12/2019].
(16) Panopto. *8 Active learning examples in higher education.* Available from: https://www.panopto.com/blog/8-active-learning-examples-in-higher-education/ [Accessed 08/12/2019].
(17) Masters JC at al. Assessment of Multiple-Choice Questions in Selected Test Banks. *Journal of Nursing Education.* 2001.
(18) Simkin MG, Kuechler WL. Multiple-Choice Tests and Student Understanding: What Is the Connection? *Decision Sciences Journal of Innovative Education.* 2004.
(19) Walstad WB, Becker WE. Achievement Differences on Multiple-Choice and Essay Tests in Economics. *The American Economic Review.* 1994.
(20) IChemE. *Accreditation of chemical engineering programmes. A guide for higher education providers and assessors.* V3.0. IChemE, 2019.
(21) Rosenthal JS. Active-Learning Strategies in Advanced Mathematics Classes. *Studies in Higher Education.* 1995223-228.
(22) Petty G. *Teaching Today: A Practical Guide*: Nelson Thornes; 2004.
(23) Zull JE. *The Art of Changing the Brain: Enriching Teaching by Exploring the Boiology of learning.* Sterling, Virginia: Stylus; 2002.
(24) Prytula M, Weiman K. Collaborative professional development: An examination of changes in teacher identity through the professional leaning community model. *Journal of Case Studies in Education.* 2012.
(25) Ericsson A, Pool R. *Peak: Secrets from the new science of expertise.* Boston: Houghton Mifflin Harcourt; 2016.
(26) Joosten T. Mobile learning and social media: Increasing engagement and interactivity. *New Media Consortium Conference, 2019*
(27) Sanchez G. Laboratorio Virtual: Herramienta pedagógica de apoyo en el proceso de enseñanza–aprendizaje. *CIMTED Corporation.* 2017.
(28) Blottnitz H..Promoting active learning in sustainable development: experiences from a 4th year chemical engineering course . *Journal of Cleaner Production.* 2006.
(29) Freeman S et al. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the USA.* 2014.
(31) Kuh GD. Promoting Student Success: What Campus Leaders Can Do. National Survey of Student Engagement. 2005.
(32) Wilkinson K. *The integration of mobile learning app-based quiz-games in higher education teaching of anatomical sciences.* DProf. Middlesex University; 2017.
(34) Lee J, Hammer J. Gamification in Education: What, How, Why Bother. *Academic exchange.* 2011.
(35) Wineberg S, Schneider J. Inverting Bloom's Taxonomy. *Education Week.* 2009.
(36) Krathwohl DR, Anderson LW. A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy. *Longman.* 2009.
(37) Clifton SL, Schriner CL. Assessing the quality of Multiple-Choice Test Items. *Nurse Educator.* 2010;35(1): 12-16.
(38) Cook BR, Babon A. Active learning through online quizzes: better learning and less (busy) work. *Journal of Geography in Higher Education.* 2016.
(39) Rosenthal M. Qualitative research methods: Why, when, and how to conduct interviews and focus groups in pharmacy research. *Current in Pharmacy Teaching and Learning.* 2016.
(40) Kosecoff J, Fink A. How to Conduct Surveys: A Step-by-step Guide. *Thousand Oaks.* 1998.
(41) Linacre JM. Investigating Rating Scale Category Utility. *Journal of Outcome Measurement.* 1999.
(42) Lietz P. Research into questionnaire design. A summary of the literature. *International Journal of Market Research.* 2009;52(2): .
(43) O'Muircheartaigh C, Krosnick JA, Helic A. *Middle Alternatives, Acquiscence, and the Quality of Questionnaire Data.* The University of Chicago: 2000.
(44) Sosniak,L., Anderson,L (eds.) *Bloom's Taxonomy.* Chicago, IL: Univ. Chicago Press; 1994.
(45) Krueger R, Casey M. A practical guide for applied research. *A practical guide for applied research.* 2000.
(46) Breen RL. A Practical Guide to Focus Group Research. *Journal of Ge.* 2007.
(47) Weitze L, Sobke H. Quizzing to Become an Engineering - A commercial Quiz App in Higher Education. : New Perspectives in Science Education; 2016.
(48) Campbell J. Data obtained during Atrivity Chemistry Trial AY 2018/19.
(49) Oulasvirta A, Rattenbury T, Ma L, Raita E. Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing.* 2012.
(50) Crowe A, Dirks C, Wenderoth MP. Biology in Bloom: Implementing Bloom's Taxonomy to Enhance Student Learning in Biology. *CBE - Life Sciences Education.* 2008.
(51) Aziz SM. Online Quizzes for Enhancing Student Learning in a First Year Engineering Course. *July 21-25.*: International Conference on Engineering Education; 2003.
(52) Astarita G. Dimensional analysis, scaling, and orders of magnitude. *Chemical Engineering Science.* 1997;52(24): .
(53) Hattie J. *Influences on student learning.* Inaugural lecture. 1999.
(54) Ahlfeldt S, Mehta S, Sellnow T. Measurement and analysis of student engagement. *Higher Education Research & Development.* 2007;24(1)

Studying the Micro-Scale Wettability of Multiphase Systems using Atomic Force Microscopy

Yash Nirmal[1], Huzaifa Patel[1]

Abstract
A procedure is outlined which characterises the wettability of crude oil drops on mica using atomic force microscopy (AFM). Two liquid-liquid-solid systems were investigated: (1) Brine-soaked calcite displaced by decane to produce water micro-droplets; (2) Oil-soaked mica imaged in water; to produce micro-droplets of crude oil. Water droplets were not produced in System (1). Crude oil droplets varying in diameter from 1 to 12 μm were observed for System (2). The images obtained gave the 3-D topography of the surface and water-oil interface, allowing contact angles for the droplets to be found via 2-D cross sectional analysis. A contact angle distribution across all the measured droplets was determined, and varied between 10.9° to 53.3°. A linear relationship was established between droplet tallness and mean contact angle. Therefore, insights were gained into the flocculation of asphaltene droplets.

Keywords
AFM — Asphaltene — Calcite — Enhanced Oil Recovery — Mica — Wettability

[1] *Department of Chemical Engineering, Imperial College of Science, Technology and Medicine, London, United Kingdom*

Introduction

When a liquid is spread over a surface, it can completely wet, partially wet or not wet the surface. The degree of contact between the liquid and solid surface is defined as wettability. However, the concepts used to define wetting are length scale dependent. At the molecular scale, wetting is defined as the adhesion between the solid and liquid phase. Surface roughness becomes an influencing factor at the nano-scale, which is when contact angles are first observed and measured. Chemical heterogeneities also become significant at this scale. Wetting is typically characterised by the contact angle, which can be determined using suitable mathematical models or experiments. The simplest of these models is Young's equation [1], which describes the equilibrium which arises from the interaction of surface tensions in an ideal gas-liquid-solid (multi-phase) system. However, real surfaces are rough and chemically heterogeneous, therefore they exhibit contact angle hysteresis [2]. This is the phenomena in which a range of contact angles are observed for a given multi-phase system. Therefore, it is imperative to experimentally measure the contact angle for most real systems.

Traditionally, contact angles have been measured using goniometry, which uses high-resolution cameras and software to accurately image droplets and determine their contact angles with a given surface. These measurements fail to capture the underlying surface structure, thus the effective wettability of the system is not captured.

Rose petals exhibit micropapillae which each have many nanofolds [3]. The "petal effect" occurs when a rose petal is turned upside down and its spherical water droplets do not fall off the surface due to their surface adhesion.

In contrast, the lotus plant has smaller micro and nano structures than the rose petal [4]. Therefore, water is not able to wet the space between the micro-structures. Hence, air remains on the lotus surface, which reduces the overall water-solid adhesive force. This results in a self cleaning phenomenon where water droplets easily roll off the lotus surface. This phenomenon shows that surface features on the micro scale can cause a radical alteration in observed wettability behaviour at the mm scale.

Oil reservoirs comprise of a multiphase rock-brine-oil mixture which exhibits complex wettability behaviour. Enhanced Oil Recovery (EOR) methods involve the injection of a fluid into reservoir rocks, to displace crude that is unobtainable from traditional oil extraction methods. Surfactant flooding achieves this by injecting dilute solutions into the reservoir, which reduce interfacial tension, thus increasing the wettability of the porous rocks. The presence of ions [5] and naturally adsorbed material [6] in oil-water-mineral systems can also induce wettability changes. Modelling the complex wettability interactions of reservoirs presents a significant economic opportunity, with associated scientific challenges. So, characterising the wetting behaviour of both crude oil and water on mineral surfaces is of key importance.

Understanding the wetting behaviour of water on mineral surfaces at the micro scale involves characterising the surface roughness. Current micro-CT studies analyse the porosity of the rock, but do not measure the surface roughness of the mineral surface. Consequently, they fail to capture information from the micro-scale, which provides insight into the effective

wettability of the rock surface.

Crude oil is a mixture of over 100,000 chemical constituents. Asphaltenes are a polar fraction of crude, commonly defined as the fraction that is soluble in toluene but insoluble in n-heptane.

Unstable mixtures of crude, indicated by a high Colloidal Instability Index (CII) can result in asphaltene precipitation on mineral surfaces [7]. This can result in impeded oil flow through reservoirs and fouling in downstream heat exchangers. FTIR spectroscopy has been used to image the carbon structures [8] and spatial distribution of these agglomerate droplets [9]. Asphaltene colloidal micelles are negatively charged [10] and will agglomerate to reduce surface contact, thus forming large droplets on mica surfaces [11].

Atomic Force Microscopy (AFM) has become the standard technology for studying surfaces at a high resolution. Its high lateral resolution allows spectacular surface images to be obtained. In particular, AFM has been used to characterise the micro-structures of mineral surfaces to understand their effective wetting behaviour. Adsorption energy is another metric which can be used to understand wettability at the nanoscale [12]. The formation and measurement of thin liquid films have been studied for a range of systems and conditions using AFM [13, 14, 15].

Here, we propose the use of AFM to characterise both the micro-structure and effective wettability of rock-oil-water systems. Micro-droplets of water and crude oil produced on mineral surfaces are imaged using AFM. The contact angles can be deduced from the obtained droplet topography and comparison with the underlying mineral micro-structure reveals the effective wettability of the system.

1. Methods

To obtain water micro-droplets, calcite was cleaved in brine and flushed with n-decane, as outlined in [16]. To obtain crude-oil droplets, crude oil was diluted in toluene and allowed to dry on a mica surface. This procedure was adapted from [11].

1.1 Solution Preparation

1.1.1 Brine Solution

A brine solution consisting of various salts was prepared with distilled water as shown in Table 1.

Table 1. Brine Solution Mixture

Species	Concentration (g/L)
$NaHCO_3$	0.18
$CaCl_2.2H_2O$	1.43
$MgCl_2.6H_2O$	10.26
$Na_2SO_4.10H_2O$	8.31
KCl	0.71
NaCl	22.74

Wettability and the contact angle are extremely sensitive to contaminants, which can easily change the surface tension. To ensure the risk of contamination was minimised, the glass storage jar used to prepare and store the brine was cleaned thoroughly. To remove all contaminants, the jar was rinsed with dilute (0.5-2%) Hellmanex solution, which significantly reduces the surface tension of water, therefore providing excellent wetting of the internal surface of the storage jar. This effectively displaces most dirt particles. Furthermore, its emulsifying and dispersing properties ensure that the displaced contaminants do not redeposit on the glass surface. The Hellmanex and displaced dirt particles were then rinsed out with deionised water (as opposed to tap water which contains depositable minerals). Isopropyl alcohol was then used to rinse any remaining fluid and dry out the jar.

1.1.2 Dilute Crude Oil Solution

The particular oil used, Crude 'D' was a heavy, viscous crude with a high asphaltene content. Toluene was used as the solvent to dilute the crude oil which was dried on the samples. To prepare a solution containing 100 mg/L of crude, 5.9 mL of the desired crude was pipetted into a 50 mL volumetric flask and the remainder of the volume made up with toluene [11]. To reach 10 mg/L, the 100 mg/L sample was diluted by a factor of 10 using a 10 mL volumetric flask and 1 mL of the 100 mg/L solution, and the remainder made up with toluene. The viscosity of the crude posed difficulties when handling, especially when attempting to pipette it into the volumetric flask. To ensure that as much of the measured volume entered the flask, the pipette tip was rinsed into the flask when filling. Once full, the flask was inverted several times to ensure the full dissolution of residue on the neck into the bulk solution.

1.2 Surface Preparation

1.2.1 Calcite

To minimise contamination to the sample, the calcite surface was freshly cleaved under brine solution. This way, the brine is able to wet the calcite surface instantly. This also prevents surface contact with the air, which contains dust particles that will easily deposit onto the calcite surface. When imaging such a contaminated sample, it will exhibit many surface impurities, which firstly makes the picture quality poor. Large calcite crystals ($CaCO_3$) were cleaved using a cleaving blade and hammer. This system is designed to cut at the precise angle (74°) required to obtain a relatively smooth surface. In practice, this is much more challenging to replicate. If the calcite was not lined up along one of the three rhombohedral cleavage planes, large steps and defects would have easily formed on the face. The brine can also cause the tip of the blade to corrode if left exposed for an extended period of time. This is especially important when cleaving calcite, as a dull or corroded blade will engender more surface imperfections in the cleaved surface. Despite this, when imaging, the optical camera could be used to locate regions which appeared flatter and did not exhibit these features.

Contaminants can enter the sample at all stages of the operational procedure. Hence to minimise this risk further

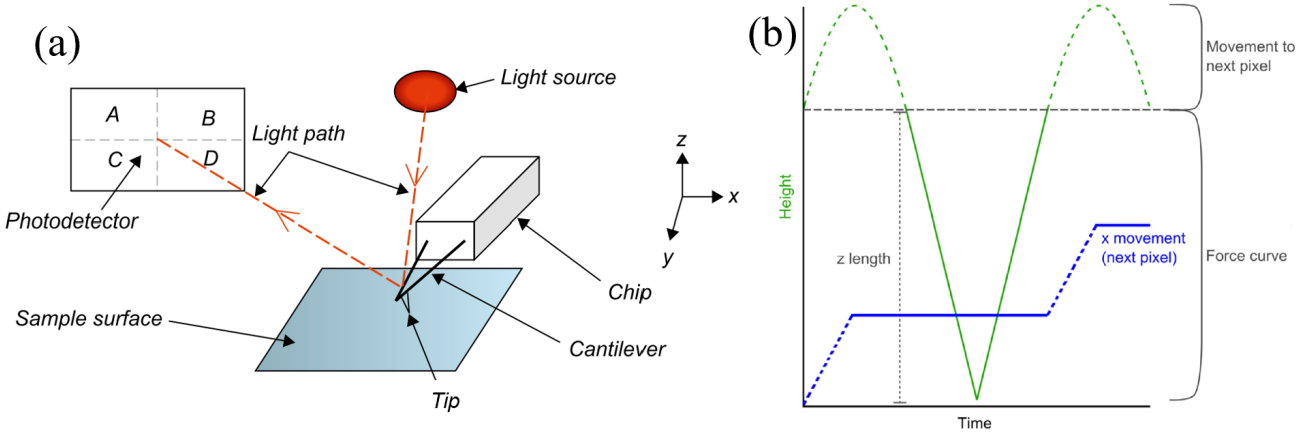


Figure 1. (a) Basic components and set-up of an AFM [17]; (b) Operation of QI mode [18]

and maintain the same basis across experiments, each sample was freshly cleaved on the day of use. When preparing a new sample, the larger crystal had to be cleaved twice to ensure that both faces of the cut sample were fresh.

1.2.2 Fluid Film Displacement

Producing micron-scale water droplets requires a small amount of wetted brine to remain on the surface, where it remains by adhesion to the surface. Too much wetted brine on the surface can easily cause the formation of a water film. Hence, the amount of brine displaced was a significant control factor and can be altered with use of different displacement methods. To displace some of the wetted brine, the calcite surface was flushed with *n*-decane [16]. Plastic droppers, plastic syringes and glass syringes with a narrow nozzle were used to flush the different samples, thus providing a range of flowrates and flowtypes which impacted the calcite surface. It could then be determined as to which displacement method was the most suitable to produce water droplets. Finally, the sample was placed in a petri dish and submerged past the top surface with *n*-decane.

1.2.3 Mica

The disk of mica was cleaved by applying and removing a piece of adhesive tape from the surface, which removed the surface layer of mica and left behind a fresh surface. The fresh mica disk was then placed in a petri dish and up to 200 μL of the dilute crude solution deposited onto the surface with a pipette, before being sealed and left to dry overnight. This resulted in a sample with dried crude on the surface which could be imaged with the AFM using water as the secondary fluid.

Crude Oil Viscous Layer Formation Initially, a concentration of 100 mg/L of crude oil in toluene solution was dried on the mica surface. When attempting to image this, it was found that the crude had formed a layer over the whole mica surface where it had been deposited. Attempts at imaging this layer yielded unconventional force-distance curves and no valid images. Upon removing and examining the AFM chip under an optical microscope, contamination was observed in the form of crude stuck to the body of the chip. This covered the cantilever and would have impeded its movement, providing an explanation as to why no proper images were obtained. It was decided to dilute the crude oil further, to 10 mg/L, in order to reduce the quantity present when depositing it onto the mica surface. After making this operational change, crude oil droplets were in fact observed.

1.3 Operation of Atomic Force Microscope

A NanoWizard 4 Nanoscience AFM designed by JPK Instruments was used to take measurements and the probes used were the 'PPP-NCHAuD-10' from NANOSENSORS. Various modes of imaging exist for AFM measurements but JPK's Quantitative Imaging (QI) was used throughout. The parameters which could be altered when imaging in QI mode were: force setpoint, z-length, z-speed and pixel time, scan size and image resolution.

1.3.1 AFM Theory

An AFM images a sample through the deflections of a cantilever with an atomically sharp tip. The tip interacts with the sample through a range of intermolecular forces. As a result, the cantilever bends and reflects a laser which shines onto a photodiode, allowing the deflections of the cantilever to be measured by the change in signal. A group of piezoelectric ceramics are used to control the position of the tip based with extremely high degrees of precision. Figure 1(a) shows a simplified diagram of how the various parts are arranged.

1.3.2 Imaging Modes

Tapping Mode Measurements taken using tapping mode involve oscillation of the cantilever at or near its resonance frequency [19]. Tip-Sample interactions as the cantilever approaches the sample change the frequency and amplitude of the oscillations. By utilising a feedback loop to keep amplitude constant, the topography of the surface is given by the height found while scanning.

Force Spectroscopy Atomic Force Microscopy can also be used to measure force-distance curves between the tip and sample for the approach and retract movements [20]. This is conventionally a time consuming measurement to carry out, and is usually taken at lower resolutions down to 16 by 16 pixels. Two examples of force spectroscopy curves obtained from measurements in QI mode can be found in Figure 4.

QI™ Mode JPK's Quantitative Imaging (QI) Mode combines the measurement speed of tapping mode with the information provided by force spectroscopy curves. Figure 1(b) contains a graphical representation of how QI mode operates. By only changing the x-position when the z-length has been reached, a full force-distance curve is measured for each pixel [18]. Since the vertical movement speed is held constant, the cantilever deflection remains unaffected and higher resolution images can be obtained at a much faster rate. The force-distance curves for each pixel open up the potential for more insightful further analysis. This mode also removes many of the complications associated with selecting the appropriate parameters for the feedback and control which would otherwise be required for tapping mode.

1.3.3 Imaging in Liquid

Most of the images taken were carried out with the sample in a liquid medium [21]. The calcite samples were immersed in *n*-decane and the mica disks had water pipetted onto the surface. As such, it was essential that there were no air bubbles present between the probe and the sample which would have interfered with the measurements. To mitigate against this, a small volume of the imaging liquid was dropped onto the AFM mount where the chip was held. When the head was placed over the sample and lowered into place, the liquid on the chip coalesced with the liquid on the sample and created a homogeneous volume without any air bubbles.

1.3.4 Imaging Parameters for QI

Scan Size and Resolution The size and resolution reflects the dimensions and detail of the square imaged. The maximum area available to image had side lengths of 100 microns. The image resolutions available were 128, 256, 512 and 1024 pixels squared. Initial images to investigate whether desirable features were present were taken at 100 microns and 128 pixels to maximise the potential of finding droplets and minimise the time taken to complete the image. When considering features in more detail, the scan size was reduced as appropriate to encompass the area of interest and the resolution was increased to 256 pixels. Increasing pixel resolution further than this was time prohibitive, since each step up increased scan time by a factor of 4 (doubling both dimensions).

Z-speed and Pixel Time These parameters share an inverse relationship: increasing the speed with which the z-movement takes place causes a proportional decrease in the time taken to carry out each individual pixel measurement and the scan as a whole. The z-speed was usually set slower for the initial image since the optimum was not known and increased when taking smaller and higher resolution images to reduce the scan time.

Force Setpoint The force setpoint determines the point at which the vertical movement will change direction: going from approaching towards the sample to retracting away from the surface. It is reached once the tip contacts the sample or interface and the cantilever bends sufficiently. The value that must be selected is highly dependent on the sample which is being measured. When considering systems with two liquid phases, the cantilever must often cross the fluid-fluid interface before contacting the solid below. In the systems studied carried out, these were decane-brine and water-crude oil.

Z-Length The z-length (shown in Figure 1(b)) of a measurement is the vertical distance cleared by the probe between individual measurements. Selecting a value is also a highly sample dependent process. For example, mica is an extremely flat surface with deviations only on the nanometre scale, so a z-length in the order of tens of nanometres is sufficient for the plain surface. Z-length and z-speed combine to give the pixel time, so an increase in z-length increases the total scan time.

1.4 Crude Oil Contact Angle Determination

To measure the contact angle of a chosen crude oil droplet, a cross section through the centre was obtained using the JPK Data Processing software. Multiple length and height measurements were taken from either end, and by applying the inverse tangent trigonometric function, the contact angle was calculated. For each droplet, eight different contact angles were measured by taking four separate cross sections along the horizontal, vertical, and diagonal directions. Each individual contact angle was calculated from the mean of three measurements taken across different spans at the end of the cross section.

2. Results & Discussion

Three-phase systems (fluid-fluid-solid) can be characterised using AFM, which makes it an ideal tool for understanding the wettability of mineral surfaces. Deflections in the adhesion can indicate the presence of a substrate on the sample surface, but can also originate from noise and vibrations. JPK's data processing software converts this height at each contacted point into a colourised image for ease of analysis. Similarly, images are obtained for the adhesion, force and stiffness of each sample. The height, adhesion and stiffness images can be compared to obtain insights into the sample surface.

Recent studies have demonstrated the use of AFM to determine contact angle [22, 23]. In this study, the approach was further characterised to study the contact angle in two systems; (1) Water droplets on a calcite surface in the presence of n-decane; (2) Crude Oil droplets on a mica surface in the presence of water. QI mode was used to obtain the force spectroscopy curve for each sampled point. The presence of water droplets can be confirmed by an inflection towards

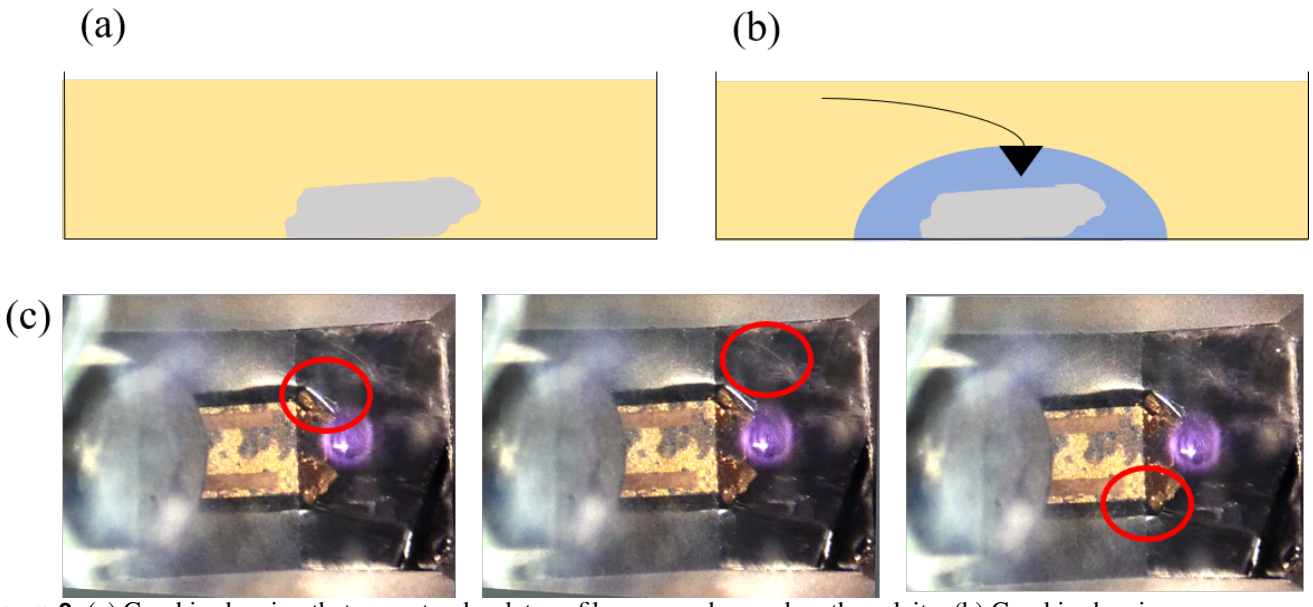


Figure 2. (a) Graphic showing that no water droplets or films were observed on the calcite; (b) Graphic showing an immeasurable thick film that formed on the calcite surface; (c) Stills from the AFM optical camera showing features of a thick film as the AFM was retracted

the surface in the force-distance profile (Figure 4(b)), which represents the tip crossing the decane-water droplet interface.

Water droplets were not observed in System (1) which is discussed further in Section 2.2. By measuring contact angles of the crude oil droplets, understanding was gained into the wettability behaviour of System (2), contained in Sections 2.3 and 2.4.

2.1 Sample Surfaces

2.1.1 Calcite

Calcite ($CaCO_3$) is a crystal polymorph of calcium carbonate. It is positively charged at neutral pH and is very prevalent in reservoir rock. Calcite can be cleaved in brine, which prevents its contact with air and ensures the formation of a liquid film. Figure 3(d) shows the three dimensional profile of the step-like kinks on the calcite surface, around the magnitude of 10 nm. These are most likely caused by improper alignment or a dull section of the blade when cleaving the crystal. To obtain a smooth cleaved surface, the crystal must be exactly aligned along one of the three cleavage planes. As the apparatus requires one to use tweezers to adjust the position of the crystal, this does not provide the full precision required to guarantee a clean cut. Flatter sections are still present on the crystal, and could be found and imaged through the assistance of the optical camera.

2.1.2 Mica

Mica are a group of phyllosilicate materials similar to clay, which are negatively charged. They exhibit basal cleavage (only on one plane) and are extremely flat at the sub-pore scale. Mica is also present in reservoir rocks, so its wetting behaviour is also relevant for industrial study. Mica was much easier to consistently cleave and obtain a flat surface, with the only deviations due to the adhesive tape not sticking properly to parts of the surface. This created areas which were still flat as a whole, but offset vertically from the rest of the surface.

2.2 Brine Droplets on Calcite Surfaces

AFM in QI mode can be used to simultaneously measure the contact angle of brine droplets and the underlying surface roughness. The robustness of the contact angle measurements are improved, since knowledge of the surface topography can explain deviations from the intrinsic contact angle. Furthermore, removing the need for goniometry to measure wettability allows more time to be spent on experimental activities, especially in time constrained experiments. With the aim of testing the reproducibility of the results found in [16], System (1) was investigated. The phenomenon of micron-scale brine droplets when following the outlined procedure had only been observed once.

Despite cleaving several calcite samples and trying three flushing methods, water micro-droplets were not observed. Figure 2(b) shows a schematic view of a thick water film which was observed on the calcite surface, but could not be measured as the thickness exceeded the AFM's maximum z-length. Figure 2(c) has three stills from the optical camera showing the AFM chip partially retracted from the calcite surface in the presence of a thick film. The first picture shows the film deforming around the top right corner of the chip, catching the light from the camera. The second picture circles the edge of the film as it pulls away from the viewing window. The third picture circles the reflection of the light catching the film at the bottom right corner of the AFM chip. Furthermore, the film could be seen with the naked eye upon the removal

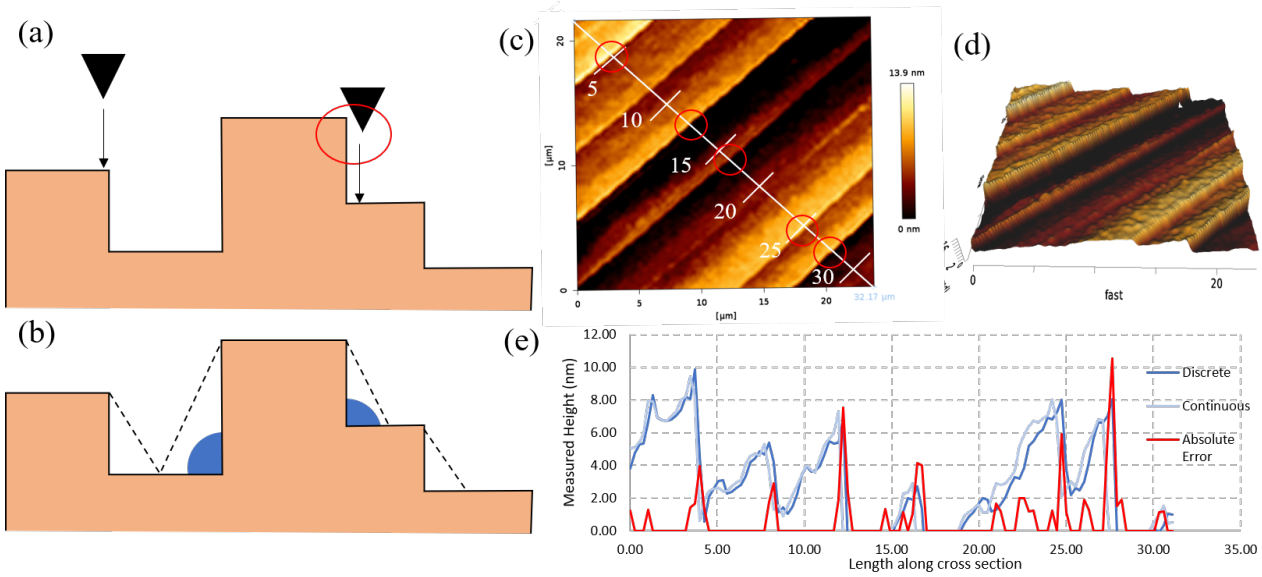


Figure 3. (a) Schematic showing AFM tip collision with steep calcite edges; (b) Schematic showing reduced imaging area due to tip collision with calcite surface features; (c) 2-D colourised height map indicating cross section diagonal and areas of imaging uncertainty; (d) 3-D colourised height map illustrating the sharp calcite steps; (e) Height of calcite surface plotted using discrete steps between AFM measurements and continuous joining of AFM points. Absolute error between the two plots demonstrates imaging uncertainty along the surface.

of the sample. To measure film thickness and the adhesive properties, the tip needs to be fully clear and disconnected from the film between force curve measurements. This gives a maximum measurable film or feature thickness as less than 6 μm, since some additional vertical clearance is required.

2.2.1 Surfactants

Contamination could have entered the process in the form of the soap used to clean the brine bottle, which may have not been fully rinsed out. Soaps work by acting as surfactants, which lower the surface tension between two liquids. Typically, this works by reducing the cohesion between water molecules, hence water droplets become weaker and break apart sooner. Therefore the surface tension of the brine droplets could have been lowered, changing the wetting behaviour. The broken droplets would instead form a continuous film along the calcite surface.

2.2.2 Limitations of AFM on Calcite Steps

Figure 3(a) and (b) show how AFM imaging can miss details in the corners when imaging steep slopes and steps. The two depictions of the tip in Figure 3(a) show the difference between impacting the corner and edge of a step. Although the second tips point has cleared the corner of the step, the side of it will still make contact with the corner. Figure 3(b) shows the profile of the steps which would actually be obtained when measuring, particularly at lower resolutions. Figure 3(c) and (d) show the 2-D and 3-D images of a calcite surface, similar to the schematics shown. Figure 3(e) graphs the height profile taken from the cross section in (c) in two different ways. It provides a quantitative estimation of the uncertainty of the surface topography across a particular cross section. The dark blue line represents the discrete height measurements obtained at each pixel along the line. The light blue line represents the imaged height when these discrete measurements are made into a continuous surface. The absolute difference between these two gives an error which is a measure of the uncertainty of the surface topography at the given point. By filtering errors less than 1 micron, this line consists of all the parts of the cross section which exhibit more significant errors. The spikes on the red line mainly correspond to large changes in height of the cross section when descending a step. It is in these steps that water is most likely to be observed, as it will adhere favourably towards both of the faces in the inside corner of the ridge. Hence it is possible that thin water films were present along the inside of the steps, but were not found in the image because the tip was unable to contact them.

2.2.3 Inconsistency in Film Production

When seeking to image the brine droplets on calcite in decane with AFM, no measurements of micron-scale water films or droplets were obtained. It was observed that fluid film formation is sensitive to: calcite cleaving, flushing method with n-decane, and sample handling in moving to the AFM. Either no water films or mm scale water films were observed, as shown in Figure 2(a) and (b). Flushing with a plastic pipette usually did not yield films, where the large nozzle area could have caused all of the water on the calcite to be flushed away by n-decane. On the other hand, flushing using a syringe needle usually resulted in mm scale films. Varying the diameter of the nozzle did not increase the precision of water film production. Given the weak correlation between

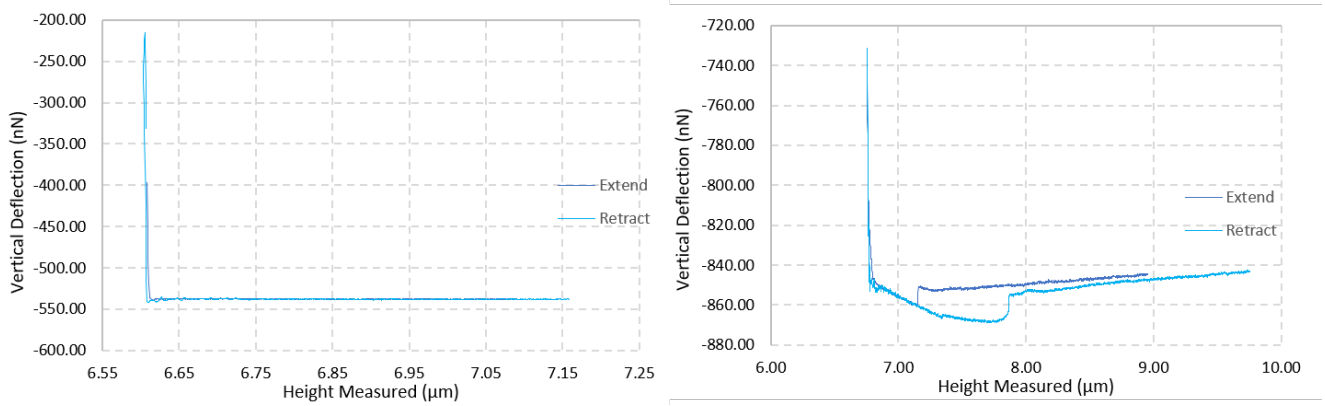


Figure 4. (a) Standard force-distance curve from calcite surface with no film observed; (b) Force-distance curve exhibiting liquid-liquid interfacial snap-in behaviour

operational parameters and the thickness of the produced film, it was concluded that the approach of flushing the brine-cleaved calcite was not precise enough to obtain micron-scale water droplets.

Force-Distance Curves for Films Where no film was observed, the force-distance curve looked like Figure 4(a). For comparison, the example curve from Figure 4(b) exhibits clear snap-in behaviour on the extend curve and adhesive behaviour on the retract curve. The film thickness is the distance between the downwards inflection and steep upwards slope increase. This represents the distance between the tip crossing the fluid-fluid interface and making contact with the solid surface below. On the retrace curve, the minimum value is the adhesive force between the tip and the film. The retrace curve cannot be used to determine the film thickness, as the water remains in contact with the tip even after the z-length clears the film. This is due to a force balance between the capillary force and the van der Waals water/surface attraction [24].

2.3 Crude Oil Droplets on Mica

It was decided that obtaining micron-scale water droplets on calcite was not reproducible using the brine cleaving and *n*-decane flushing procedure. Furthermore, producing brine microdroplets using a different operational procedure was deemed unlikely since film formation was erratic and sensitive to operational parameters. Hence, the decision was made to shift the focus onto studying the wettability of crude oil droplets on mica surfaces.

Most studies into asphaltene deposition use n-heptane as a solvent, however it has been shown that a Smoluchowski model can effectively represent a toluene-dissolved flocculation of asphaltenes [25]. The wettability behaviour of large crude oil 'D' deposits on mica was understood, by obtaining multiple contact angles for each droplet. Crude oil does not form droplets on calcite surfaces, due to the opposing charge (asphaltene micelles are negative and calcite is positively charged). Therefore, using a flat mica surface negated the influence of surface roughness on wettability, hence the effect of chemical heterogeneities could be singled out.

2.3.1 Crude Oil Droplets

After the 10 mg/L of crude oil dissolved in toluene was left to dry overnight on the mica sample, several crude oil microdroplets were observed the next day. These were imaged in water with a pseudo-tapping approach using the AFM. A low setpoint of 50 nN was chosen such that it would be reached upon contact with the crude oil at the fluid-fluid interface. This had further benefits of preventing both deformation of the droplet profile and contamination on the probe as was observed at the higher concentration (Section 1.1.2). The 3-D topography of the crude-water interface was measured, enabling 2-D cross sections to be taken in various orientations around the droplet and the respective contact angles measured. However, this approach did not allow the underlying surface structure to be imaged. The crude droplet was not fully penetrated and the height measurement of the mica surface below was not obtained. This did not have any ramifications for the contact angle analysis because cleaved mica is extremely flat and the droplet formation would not have been adversely affected.

2.4 Analysis of Crude Oil Contact Angles

Figure 5(a) shows a variety of crude oil droplets on mica, which vary in shape and diameter (1-12 μm). Two distinct types of droplets were observed. Smaller droplets tend to be more spherically shaped, whereas larger droplets tend to agglomerate together and often show distinct peaks. One such droplet can be clearly seen in the bottom right of Figure 5(a). These larger, multiple peaked droplets have to balance the effect of minimising surface contact with mica and holding up their own weight. To study the wettability of these droplets, contact angles were measured using the droplet topography obtained. Contact angles were determined from four cross sections taken across the diameter of each droplet as depicted in Figure 5(a) by the green lines. Figure 5(b) and (c) shows the two marked crude oil droplets from Figure 5(a), with their circumferential contact angles and the average of all eight values.

Figure 6 (a) shows that the droplets measured exhibited

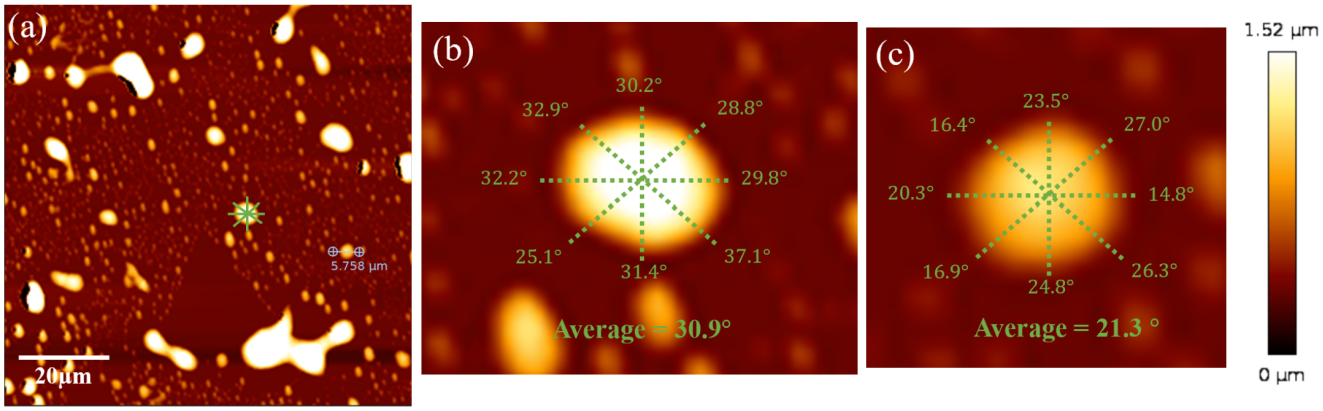


Figure 5. (a) 90 micron square image of crude oil droplets on mica surface, with cross section axes and sample cross section line overlaid; (b) Close-up of droplet with green axes from first image; (c) Close-up of droplet with blue software overlay from first image [Scale Bar applicable to all images]

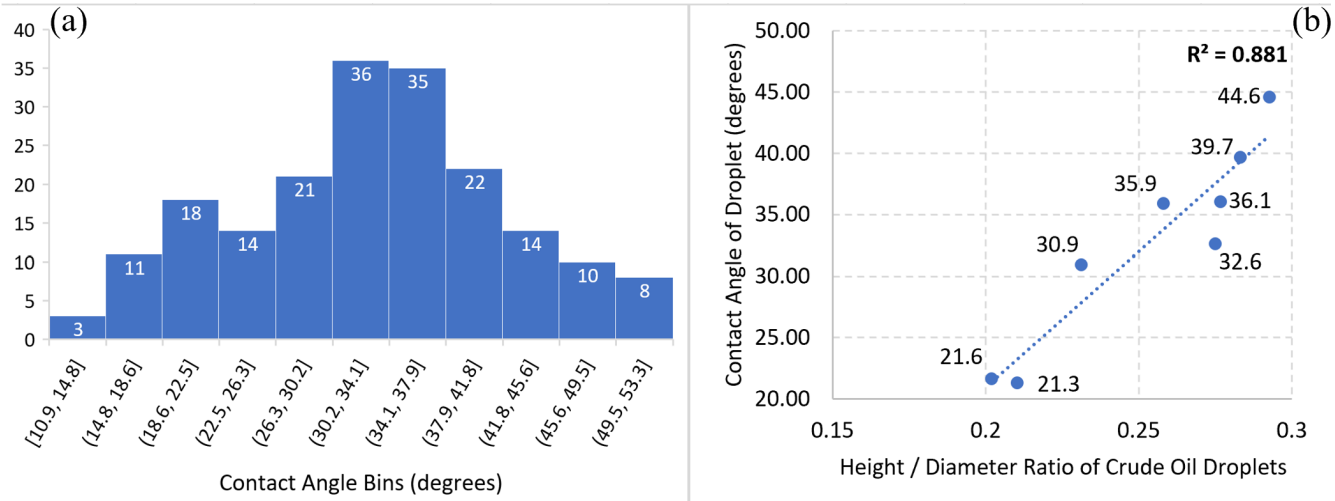


Figure 6. (a) Correlation between 'tallness' of droplet and mean contact angle measured; (b) Histogram containing the distribution of all crude oil contact angles measured

a large overall contact angle variance. Upon observation of the droplets (b) and (c) in Figure 5, it can clearly be seen that there exists no clear trend to the contact angle behaviour. The maximum and minimum values for droplet 5(c) are observed adjacent to each other. Therefore, the asphaltene agglomerates do not exhibit typical droplet behaviour. However, it is more likely that the variance of contact angles resulted from the relatively low resolution (256x256) of the image and the small portion of this that the droplet occupied. Due to the low resolution, not enough height points were sampled at the edge of the droplet. To reduce the effect of this variation in further analysis, the mean contact angle was taken for each droplet.

The droplet in Figure 5 (b) exhibits a much higher mean contact angle of 30.9° compared to 21.3° for the droplet 5(c). It can be seen from the colourised height image that droplet 5(b) is much taller than droplet 5(c). More asphaltene micelles agglomerate for droplet (b) and to minimise the surface contact with mica, they form taller droplets, resulting in a larger contact angle.

To gain a deeper understanding of asphaltene flocculation, the mean contact angles of several droplets were compared to the droplet 'tallness', which was defined as follows:

$$\textit{Droplet Tallness} = \textit{Droplet Height} \ / \ \textit{Droplet Diameter}$$

Figure 6(b) shows an $R^2 = 0.88$, thus indicating a fairly strong correlation between contact angle and droplet tallness. This trend is applicable to single droplets which have a circular circumferential profile, as the energetics of larger combined droplets differ.

3. Conclusion

The procedure followed to obtain the brine microdroplets was unable to be reproduced. Two of the stages involved were identified as critical for producing droplets. Cleavage of the calcite crystal required a high degree of precision, which required skill to obtain the exact alignment. When done incorrectly, surface features like steps and ridges were formed, hindering the production and observation of droplets. Displacement of

the fluid film was also inconsistent, especially when using a plastic dropper or plastic syringe as the flushing tool. A thick fluid film was observed when using a glass syringe with a thinner nozzle, which provided a focused stream of decane at a consistent flowrate. This fluid film had a thickness on the millimetre scale, as it was observed through the optical camera mounted to the AFM. It could not be imaged because the probe had insufficient vertical clearance to disconnect from the brine-decane interface.

AFM was applied to measure the contact angles of crude oil droplets on a mica surface in water. It was shown that the distribution of these angles showed a large variation, ranging from a minimum of 10.9° to a maximum of 53.3°. The mean contact angle for each droplet was plotted against the droplet 'tallness', and a linear correlation with an R^2 of 0.88 found. This provides evidence for the intuitive relationship that comes to mind when picturing the effects of droplet height and diameter. Furthermore, there is a theoretical basis for this relation: both the mica surface and asphaltene micelles are negatively charged, and the contact angle increases to minimise surface contact for larger droplets.

4. Outlook

4.1 Brine Droplets on Calcite

As it was concluded that the brine cleaving, n-decane flushing procedure was not reproducible, several suggestions are made for improving the procedure to increase both the sample quality and the probability of forming brine droplets. It is recommended that the calcite cleaver be modified to include an element which serves to align the calcite sample parallel to the cleaving blade. This can reduce surface imperfections in the calcite, thus improving the chances of water droplets being imaged. Additionally, noise from impurities can be reduced with the implementation of a rigorous cleaning procedure. This could include the UV treatment and DI rinsing of tips and thorough rinsing of storage jars and glassware to remove surfactants. Moreover, molecular sieves could be placed in the bulk solution of the petri dish to draw water out of mm-scale films and obtain thinner imageable films or brine droplets.

4.2 Crude Oil Droplets on Mica

Given the wettability insights obtained from crude oil droplet experiments, recommendations are made to improve the imaging procedure and gain deeper insights from measurements. To obtain valid contact angle distributions, multiple droplets should be imaged to increase the sample size. Therefore, a relatively large image (above $30\mu m$) should be taken. To obtain accurate contact angles for images of this scale, the imaging resolution should be set to at least 512x512. In contrast, to quickly characterise the crude oil wettability of a given surface, the contact angles of a single droplet can be accurately measured using a resolution of 128x128. To further characterise the nature of asphaltene flocculation on mica, classification methods can be applied to study the size and nature of droplets. Furthermore, droplet topographies after given drying times could be compared to predictions from a Smoluchowski model. Several crudes can be compared in varying concentrations, to reveal the differences in asphaltene droplet structure and shape.

References

[1] R. Tadmor, "Line energy and the relation between advancing, receding, and Young contact angles," *Langmuir*, vol. 20, no. 18, pp. 7659–7664, 2004.

[2] D. Quéré, "Wetting and Roughness," *Annual Review of Materials Research*, vol. 38, pp. 71–99, aug 2008.

[3] L. Feng, Y. Zhang, J. Xi, Y. Zhu, N. Wang, F. Xia, and L. Jiang, "Petal effect: A superhydrophobic state with high adhesive force," *Langmuir*, vol. 24, no. 8, pp. 4114–4119, 2008.

[4] A. Marmur, "The lotus effect: Superhydrophobicity and metastability," *Langmuir*, vol. 20, no. 9, pp. 3517–3519, 2004.

[5] F. Mugele, B. Bera, A. Cavalli, I. Siretanu, A. Maestro, M. Duits, M. Cohen-Stuart, D. Van Den Ende, I. Stocker, and I. Collins, "Ion adsorption-induced wetting transition in oil-water-mineral systems," *Scientific Reports*, vol. 5, pp. 1–8, 2015.

[6] J. Matthiesen, N. Bovet, E. Hilner, M. P. Andersson, D. A. Schmidt, K. J. Webb, K. N. Dalby, T. Hassenkam, J. Crouch, I. R. Collins, and S. L. Stipp, "How naturally adsorbed material on minerals affects low salinity enhanced oil recovery," *Energy and Fuels*, vol. 28, no. 8, pp. 4849–4858, 2014.

[7] S. Ashoori, M. Sharifi, M. Masoumi, and M. Mohammad Salehi, "The relationship between SARA fractions and crude oil stability," *Egyptian Journal of Petroleum*, vol. 26, no. 1, pp. 209–213, 2017.

[8] F. H. Tay and S. G. Kazarian, "Investigating the Carbon Structures of Crude Oil Deposits and Asphaltenes Combining Raman and ATR-FTIR Spectroscopy," in *AIP Conference Proceedings*, vol. 1267, pp. 457–458, 2010.

[9] A. S. Shalygin, I. V. Kozhevnikov, S. G. Kazarian, and O. N. Martyanov, "Spectroscopic imaging of deposition of asphaltenes from crude oil under flow," *Journal of Petroleum Science and Engineering*, vol. 181, no. June, p. 106205, 2019.

[10] S. E. Taylor, "The electrodeposition of asphaltenes and implications for asphaltene structure and stability in crude and residual oils," *Fuel*, vol. 77, pp. 821–828, jun 1998.

[11] M. Keshavarzi, "Application of Atomic Force Microscopy to Enhanced Oil Recovery," Tech. Rep. June, Imperial College London, 2019.

[12] R. Giro, P. W. Bryant, M. Engel, R. F. Neumann, and M. B. Steiner, "Adsorption energy as a metric for wettability at the nanoscale," *Scientific Reports*, vol. 7, pp. 1–8, 2017.

[13] S. Santos and A. Verdaguer, "Imaging Water Thin Films in Ambient Conditions Using Atomic Force Microscopy," *Materials*, vol. 9, p. 182, mar 2016.

[14] M. L. Forcada, M. M. Jakas, and A. Gras-Martí, "On liquid-film thickness measurements with the atomic-force microscope," *The Journal of Chemical Physics*, vol. 95, pp. 706–708, jul 1991.

[15] E. Bonaccurso, M. Kappl, and H. J. Butt, "Thin liquid films studied by atomic force microscopy," *Current Opinion in Colloid and Interface Science*, vol. 13, pp. 107–119, jun 2008.

[16] M. Rücker, *Wettability and wettability alteration at the pore- and nano- scales*. PhD thesis, Imperial College London, 2018.

[17] D. Johnson, N. Hilal, and W. R. Bowen, "Basic Principles of Atomic Force Microscopy," in *Atomic Force Microscopy in Process Engineering*, pp. 1–30, Elsevier, 2009.

[18] JPK Instruments, "QI™ mode - Quantitative Imaging with the NanoWizard ® 3 AFM," 2011.

[19] H. Holscher, *AFM, Tapping Mode*, p. 99. Dordrecht: Springer Netherlands, 2012.

[20] F. A. Carvalho and N. C. Santos, "Atomic force microscopy-based force spectroscopy - Biological and biomedical applications," *IUBMB Life*, vol. 64, no. 6, pp. 465–472, 2012.

[21] B. W. Hoogenboom, *AFM in Liquids*, pp. 83–89. Dordrecht: Springer Netherlands, 2012.

[22] Y. Deng, L. Xu, H. Lu, H. Wang, and Y. Shi, "Direct measurement of the contact angle of water droplet on quartz in a reservoir rock with atomic force microscopy," *Chemical Engineering Science*, vol. 177, pp. 445–454, feb 2018.

[23] M. Rücker, W. B. Bartels, and K. Singh, "The Effect of Mixed Wettability on Pore - Scale Flow Regimes Based on a Flooding Experiment in Ketton Limestone Geophysical Research Letters," pp. 3225–3234, 2019.

[24] R. Ledesma-Alonso, D. Legendre, and P. Tordjeman, "AFM tip effect on a thin liquid film," *Langmuir*, vol. 29, no. 25, pp. 7749–7757, 2013.

[25] M. H. Rad, M. Tavakolian, I. Najafi, M. H. Ghazanfari, V. Taghikhania, and M. Amani, "Modeling the kinetics of asphaltene flocculation in toluene-pentane systems for the case of sonicated crude oils," *Scientia Iranica*, vol. 20, no. 3, pp. 611–616, 2013.

Investigating the potential of porous ceramic membranes prepared via the adapted freezing technique

Muhammad Adam, Haoran Xiong
Department of Chemical Engineering, Imperial College London, U.K.

Abstract In this research, porous ceramic membranes were synthesised using the phase inversion method and an adapted freezing technique to investigate there potential. The adapted freezing technique adopts the foundation of the freeze-drying method, replacing the primary drying step involving sublimation under a vacuum with an icy-water bath to remove the frozen solvent in a more energy efficient process. The morphology and performance of the porous ceramic membranes obtained from the phase inversion and freezing methods were analysed and compared with each other. The freezing membranes displayed a unique dendritic structure with micro-channels of a random orientation. The membranes showed an asymmetric structure with no selective layer. This structure was responsible for the improved water and nitrogen permeation flux observed when compared to membranes produced via phase inversion. The unique morphology and larger pore size, around 1200 nm, opens the potential for novel applications, such as catalytic membrane reactors.

1. Introduction

Membranes are becoming an increasingly popular technology through which separation is achieved across many industries and is already widely used for several gas and liquid separations [1]. It has many advantages compared to conventional methods, including the reduced use of chemicals and being more energy efficient, both leading to a more environmentally friendly process [2]. They are also relatively easy to scale up and are simple to operate [2]. Membranes have the potential to reach an exegetic efficiency of 90% [3]. This value is much higher than that of the most commonly used separation processes currently, distillation, which only has an exegetic efficiency of 10% [3]. The membrane technology industry promises huge economic potential. It is expected to rise above USD 20 billion by 2025, with a compound annual growth rate of 6.4% from 2018 to 2025 [4]. These factors highlight the importance of investigating membranes and their potential applications.

Membranes can be split into two broad categories; polymeric and ceramic. Polymeric membranes consist of organic materials whilst ceramic membranes are artificial and consist of inorganic materials. Polymeric membranes are traditionally cheaper, however they aren't able to operate in more aggressive media compared to ceramic membranes, when temperatures are high or chemical conditions are extreme for example [5]. Ceramic membranes generally have improved durability and last longer than polymeric membranes [5]. They usually have a higher permeability than polymeric membranes which leads to increased flux [5]. Membranes can also be characterised as either porous or non-porous and can exist in a variety of unique configurations; flat sheet, tubular, hollow fibre or spiral wound [5]. Porous membranes offer improved flux compared to non-porous membranes, due to the presence of a thin layer of semi-permeable material that facilitates the separation, [6] and can be produced using many materials, including ceramics, glass and metal [7]. Meanwhile, non-porous membranes rely on diffusion of the permeate through a dense film. Separation using porous membranes is dependent upon the solute molecular size and membrane pore distribution [8]. The properties of each configuration allow for unique applications.

There are several processes that presently adopt membrane technology, including microfiltration, ultrafiltration and nanofiltration. Micro and ultrafiltration are utilised when the separation consists of larger particles and favours porous membranes [6]. Nanofiltration allows for the separation of salts in water using non-porous membranes and hence is adopted for water purification [6].

The membrane itself allows selected materials to pass through, whilst others remain on the surface. The driving force of this separation technique arises through adapting various different parameters; applying high pressure, electricity or a concentration gradient.

Currently, membrane technology has a wide variety of applications. They are used within CO_2 capture processes to separate CO_2 from exhaust fumes in a process [9]. Other applications include energy conversion in flow batteries via ion separation. One of the most common and widely adopted applications for membranes is related to water purification. Membranes are highly effective at separating impurities to produce clean water [5].

2. Background

Porous ceramics are unique open-cell structures that allow gas or liquid to selectively flow through [10]. They have the potential to replace plastic or fibre-based components due to their increased performance [10].

The synthesis of porous membranes currently has several routes:

1. Sintering process: A powder of specific particle size is pressed into a film and sintered below the materials melting point. The pore size is influenced by the powder's particle size. [7]
2. Stretching technique: Polymer powder is removed at temperatures around its melting point, followed by annealing and cooling. The film is then fractured by stretching and uniform pores are produced. [7]
3. Track etching: A damaged track is formed across a non-porous polymer film by irradiating it with heavy ions. This is followed by an immersion step in acid or alkali, etching away at the polymer to produce cylindrical pores. [7]
4. Template leaching: A homogenous film is prepared using membrane material and a leachable substance, which is then removed by chemical treatment. [7]
5. Phase Inversion: One of the most established and commonly used methods is the phase inversion technique. This method involves the removal of a solvent from a liquid-polymer solution to produce a porous, solid membrane. There are four different routes that utilise the phase inversion technique; immersion precipitation, thermal-induced phase separation, precipitation from the vapour phase and precipitation by controlled evaporation. In this study, the immersion precipitation technique was adopted when producing membranes via the phase inversion method. This technique involves the casting of a polymer solution onto a support, followed by immersion into a coagulation bath containing a non-solvent. Here, the membrane is formed through solvent-non-solvent exchange [7]. The process can be tailored to produce porous membranes with a variety of pore sizes; changing the type of polymer, polymer concentration or precipitation temperature [7].

Freezing techniques are widely researched for the synthesis of porous ceramics. Different techniques offer membranes with unique morphologies due to the way they are synthesised. One of the more popular freezing techniques is the freeze-drying method. This method involves dispersing the chosen material in a fluid, usually water, followed by a freezing process which produces a pore template for the material matrix by forming a continuous crystal network of the fluid [11]. Membranes produced via freeze-drying are capable of maintaining the pore structure due to the low surface tension involved in the process [11]. Freeze-drying offers many advantages; small sintering shrinkage, simple sintering control, widely controllable porosity, and environmental friendliness [12]. However, a major setback with this method is the primary drying step, which involves placing the suspension in a freeze dryer to remove the frozen solvent by sublimation under a vacuum. This stage is highly energy intensive, hence our method investigates an alternative drying technique to reduce energy consumption within this method. The adaptation involves placing the frozen sample in an icy-water bath to remove

the frozen solvent. Research into adapting the freeze-drying process already exists; (Araki and Halloran, 2005)[13] investigates a new freeze-casting technique capable of manufacturing near room temperature to eliminate the costs associated with operating at sub-zero temperatures. This was successfully achieved by utilising an alternative vehicle to water, camphene.

The freezing technique produces porous membranes with unique structures. This arises due to the growth of ice crystals during the freezing stage. The size of the crystallites formed depends upon the kinetics of the crystallisation during the cooling stage. The cooling rate has a significant impact on the size of the crystallites formed; faster cooling rates lead to a reduced crystallite size with a narrower distribution [14]. The polymer solution is unidirectionally cooled to create a temperature gradient. At the colder end, the solution crystallises, resulting in the formation of a concentration gradient of the solute. The phenomenon Mullins-Sekerka instability can be used to explain the ice cell growth; the solute concentration increases ahead of the ice front which reduces the melting point of the solution. This forms a constitutional supercooling zone, breaking down the planar interface [11].

When investigating the morphology of membranes produced via our adapted freezing technique, we also carried out experiments to produce membranes via the phase inversion method, allowing for an easy comparison of the properties between the two. N-methyl-2-pyrrolidone (NMP) is one of the most commonly used solvents for the preparation of ceramic suspensions due to its high solvency for a wide range of organic and inorganic compounds. It is also highly miscible with water at all temperatures and hence was used when performing the phase inversion method. However, due to the nature of the freezing technique, it was not possible to use NMP and therefore an alternative solvent, dimethylsulfoxide (DMSO), was selected. This is because DMSO has a much higher freezing point (19˚C) compared to NMP (-24.2˚C), which leads to a greater temperature gradient. This superior temperature gradient is required to drive the freezing step.

3. Experimental

3.1. Methods

3.1.1. Membrane preparation – freezing

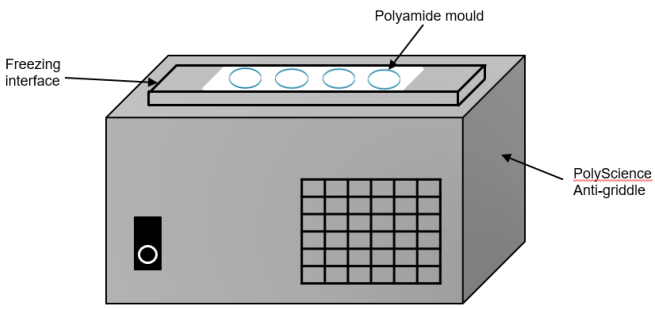


Figure 1: Schematic of the freezing apparatus; showing the PolyScience Anti-griddle equipment and polyamide mould, mounted upon the stainless-steel freezing interface

The membrane discs for the freezing technique were prepared using the following materials: 45 g of yttria-stabilized zirconia (YSZ) powder (10-100 nm, 3 mol% yttria), purchased from Inframat Advanced Materials LLC and utilised as supplied [15]. 10 g of polyethersulfone (Radal A300, Ameco Performance, USA), and 1.5 g of polyethylene glycol 30-dipolyhydroxystearate, (Arlacel P135, Uniqema) were used as the binder and dispersant respectively. Finally, 70 g of DMSO (HPLC grade, VWR) was used as the solvent. These materials were blended via ball milling to form a ceramic suspension with a high homogeneity and stability. The suspension was then poured into an isolated reservoir and degassed under a vacuum until bubbles were no longer visible. After the degassing stage, the resultant suspension was transferred to cylindrical polyamide moulds of 5 mm thickness and 40 mm diameter, which were glued onto a stainless-steel plate using silicone gel. The casting stainless-steel plate was then attached onto a pre-cooled cold plate (-30 ºC, detected by an infrared thermometer) for 30 minutes in order to freeze the discs entirely, as shown in Figure 1. The top of the mould was open, so the upper surface of suspension was exposed to the air at room temperature, 25˚C. After cooling, the membrane precursors were immersed in an iced water bath to leach the frozen solvent out for 24 hours. Finally, the precursors were sintered and calcined in air (CARBOLITE furnace) to form flat membrane discs

at a high temperature (around 1100-1200˚C), with specified temperature profile settings [13]. Initially, the temperature was increased to 600˚C from room temperature, at a rate of 2˚C per minute and held for a period of 2 hours. Following this, it was further increased to 1350˚C, at a rate of 5˚C per minute and held for a duration of 4 hours. Finally, the temperature was reduced back to room temperature at a rate of 5˚C per minute [16].

3.1.1. Membrane preparation – phase inversion

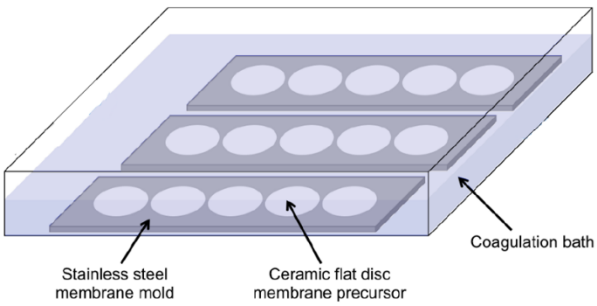


Figure 2: Schematic of the ceramic disc membrane preparation set-up for the phase inversion method [5]

The phase inversion membrane discs have a different composition compared to freezing for the suspension: 82 g of YSZ powder, 2.5 g of dispersant and 20 g of Poly-methyl 2-methylpropenoate (PMMA, Alfa Aesar) were dispersed into 100 g of NMP (HPLC grade, VWR International) to form a homogenous suspension after ball grinding. After using the same degassing technique mentioned previously, the continuous suspension was poured into a stainless-steel mould of 1 mm thickness and 30 mm diameter and placed horizontally into a coagulation of deionized water for 24 hours as shown in Figure 2 [5]. Finally, the phase inversion membrane precursors utilised similar sintering conditions and temperature profile settings in order to obtain complete ceramic membrane discs.

3.2. Membrane Characterisation

3.2.1. Water permeation flux

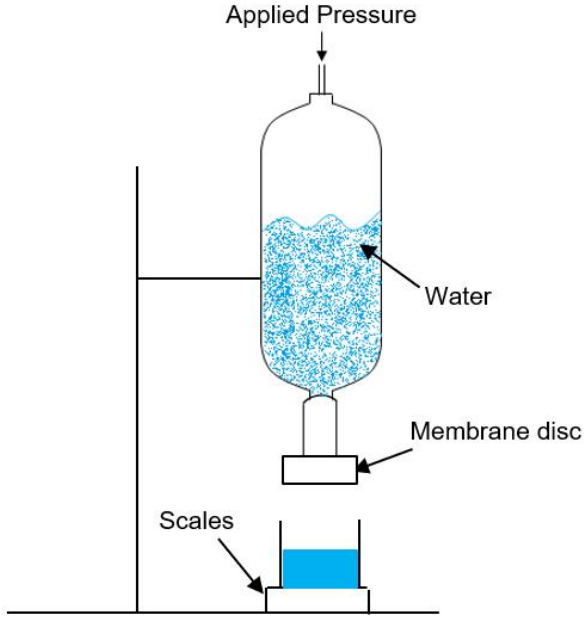


Figure 3: Schematic of the home designed dead-end system. Gas is pumped into a vessel containing water from the top, with the membrane disc at the bottom. The water permeated through is collected in a beaker at the bottom on scales

The water permeation flux was determined using a home designed dead-end system as shown in Figure 3. Tests were performed by sealing a flat disc membrane to one side of a stainless-steel tube of 17 mm internal diameter and 80 mm length with epoxy resin. To allow the water to permeate through, a constant trans-membrane pressure of 1 bar was applied. The water was collected, and its weight recorded every 5 seconds for a period of 15 minutes [5]. This was repeated three times to improve the reliability of the results. The permeation flux was then calculated based on following equation [17]:

$$J = \frac{V}{A \times t \times p} \quad (1)$$

where J is the water permeation flux (L m^{-2} hr^{-1} bar^{-1}); V is the volume of water (L); A is the effective membrane area (m^2); t is the time (hr); P is the pressure difference (bar).

3.2.2. Nitrogen permeation flux & pore size

To determine the nitrogen permeation flux and pore size, gas-liquid displacement measurements were taken using the PoroLux 1000

Porometer equipment. Epoxy resin was used as an adhesive to fit the samples, with a maximum height of 3 cm, into the sample holders. The sample was then wetted with the deionized water, filling up all the pores in the sample. Nitrogen gas between 0-4 bar with multiple steps was used to build up pressure within the system. Flow meters within the apparatus monitor the flow of nitrogen through the sample which pushes the wetting liquid out of the pores. A graph is produced using the data obtained showing the relationship between the nitrogen flow rate and pressure [18]. The pore size corresponding to each operational pressure can be obtained by using the Young-Laplace equation [17]:

$$d = \frac{4\gamma cos\theta}{\Delta P} \quad (2)$$

where d is the diameter of the pore (nm); γ is the surface tension of deionized water (72 mN/m); θ is the contacting angle of wetting liquid on the membrane surface (0˚), ΔP is the pressure difference across the membrane (bar).

3.2.3. Porosity

To calculate the porosity, a disc sample was taken, and its weight and volume measured using scales and a Vernier calliper. The disc was then inserted into water under a vacuum for one hour, followed by a second weight measurement. Finally, the disk was placed into an oven at 100 ˚C for half an hour and a final weight measurement was taken. This allowed for the porosity of the disk to be determined and was repeated a total of three times to produce reliable results.

3.2.4. Morphology

To analyse the morphology of the membranes produced, a scanning electron microscope (SEM, LEO Gemini 1525 FEGSEM, Tokyo, Japan) was used to produce high-resolution images of the membrane structure under vacuum at varying magnifications. The membrane discs were snapped in half to allow analysis of the cross-sectional structure.

4. Results and Discussion

4.1. Morphology

4.1.2. Membrane structure – phase inversion

The SEM images shown in Figure 4 are taken from (M. Lee, 2016) and display the structure of two membranes produced via the phase inversion technique, similar to the one used in this paper. Image (a) uses DMSO as the solvent, whilst image (b) uses NMP. Although the micro-channels display different shapes, they are still unidirectional regardless of the solvent used. The membrane image on the left, image (a), consists of thin, sponge-like layers, sandwiching densely packed, long and narrow micro-channels. These micro-channels initiate at the top of the membrane, propagating downwards and stopping just before the bottom of the membrane [16]. The micro channels for the image on the right, image (b), have a different orientation and show a pear-shaped structure. The micro-channels increase in size horizontally from top to bottom due to the effect of the bottom surface of the mould. The two images show very similar features. The micro-channels in both images show a uniform distribution parallel to the surface of the membrane. The length and size of the micro-channels increases from top to bottom. Towards the top, the structure is much denser, allowing for the micro-channels to be distinguished into different levels by length, size and density [16].

4.1.2. Membrane Structure – freezing

Figure 5 shows some of the SEM images taken from the membranes produced via the freezing technique with a solid loading of 35 wt%. Image (a) shows a cross-sectional image of the freezing disc at a magnification of 70 times. The membrane shows a very dense collection of micro-channels with an asymmetric structure and no selective layer. The micro-channels also display a certain level of homogeneity, with ceramic tips that are not unidirectional. This is not typical for membranes that undergo unidirectional cooling, as they show unidirectional micro-channels with an increasing diam-

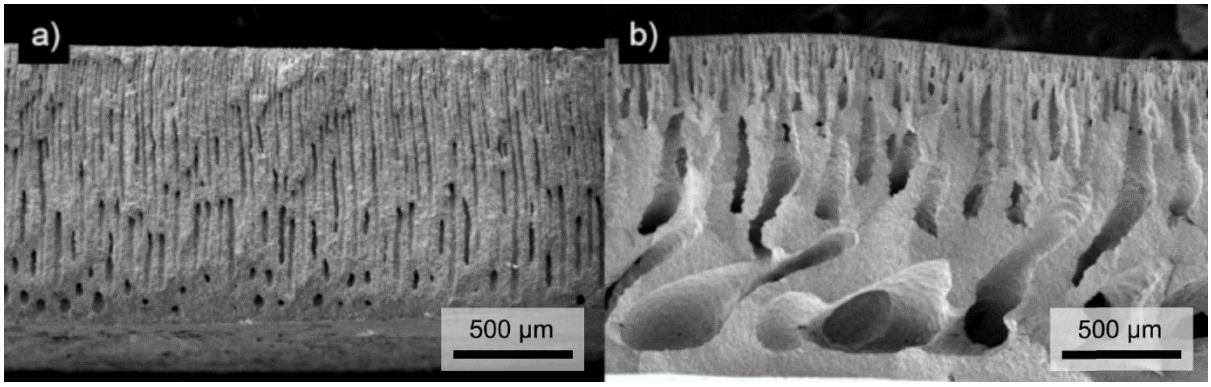


Figure 4: Cross-sectional SEM images of the upper region of: a) membrane-DMSO; b) membrane-NMP taken from (M. Lee, 2016)

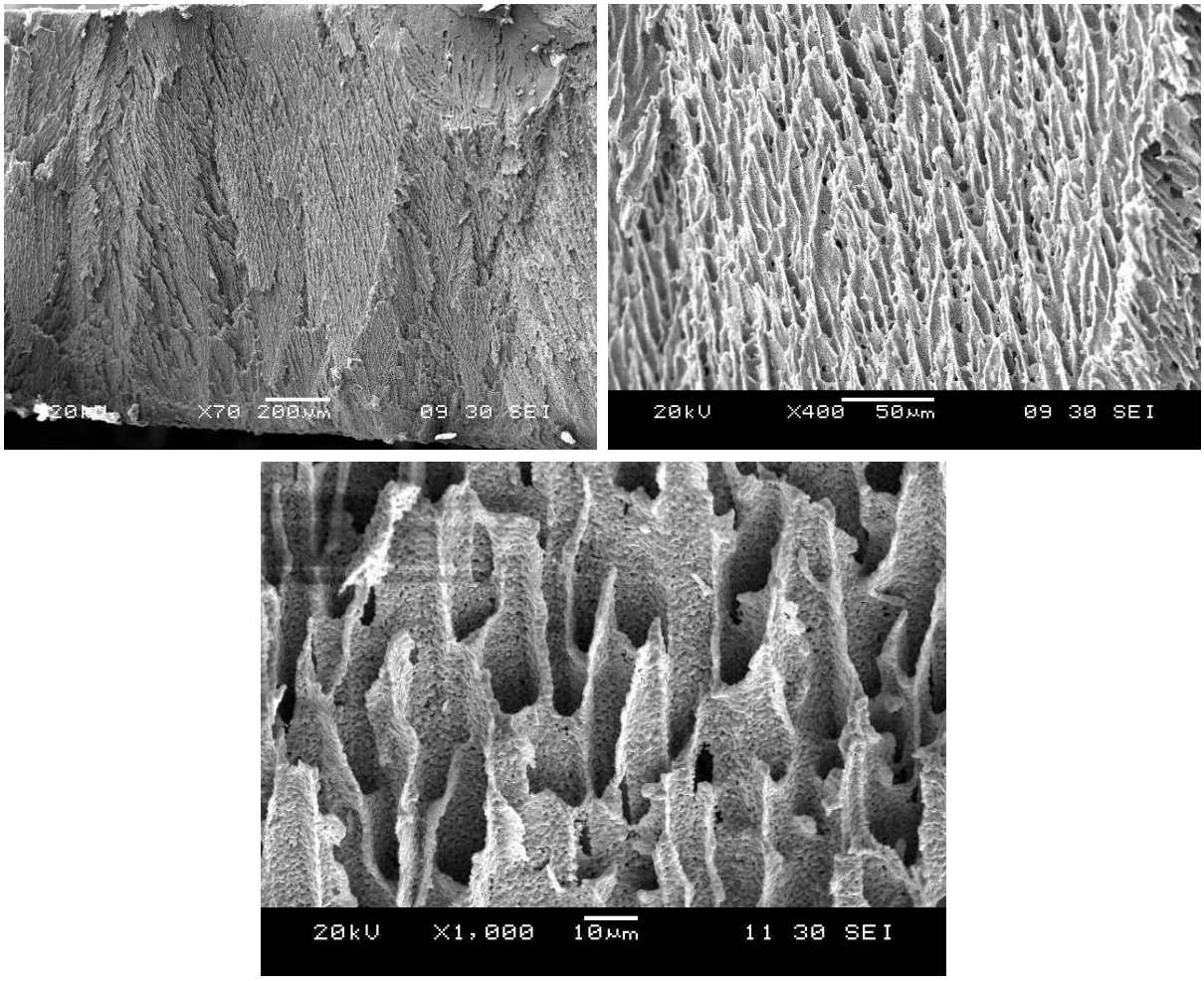


Figure 5: Cross-section SEM images of the membranes producing via the freezing technique at 35% solid loading, sintered at 1,200 ˚C. (**a**) x70 magnification; (**b**) x400 magnification; (**c**) x1,000 magnification

eter from the cold side to the top. Image (b) is the same cross-section of the membrane, further zoomed in at a magnification of 400 times and shows the consistency of the micro-channel direction in certain regions of the membrane. However, the micro-channel size is not consistent and varies throughout the image. This image further highlights the lack of unidirectional micro-channels, since the growth of the channels are not perpendicular to the surface. The image also shows the dendritic and columnar ceramic wall of the membrane, which results from the flat temperature gradient [17]. When compared to the sponge-like structure of

phase inversion membranes, the dendritic nature of the freezing membranes leads to a reduced mechanical strength. This was physically observed when handling the two membranes during the experiments and highlights the membranes brittleness. The same cross-section is further zoomed in at a magnification of 1000 times and highlights the details of the membrane produced, as shown in image (c). The ceramic displays a tip structure [17] which is due to the freezing step involved in the process. The micro-channels observed in the image show a variety of sizes. The presence of pores within the ceramic wall can also be identified, allowing for the transfer of material between micro-channels. The morphology of the membrane obtained from the freezing technique is clearly very different from that produced via phase inversion. The phase inversion membranes show unidirectional micro-channels with a consistent profile parallel to the surface. On the other hand, freezing membranes display micro-channels that are more random, with certain regions of uniformity. The sponge-like cellular structure of phase inversion membranes contributes to the increased mechanical strength compared to the ceramic tip structure of the freezing membranes, which are much more brittle. This reinforces the unique structure of the membranes produced via freezing.

4.2 Pore Size

Figure 6 shows a graph of nitrogen flowrate vs pressure at two different conditions using data obtained through the Porolux 1000 Porometer. The dry curve is interpolated to correspond to the data points for the wet curve. The first condition is when the ceramic disc was immersed into the wetting liquid and hence is represented by the wet curve on the graph. The second condition represents the case where all the liquid has been removed from the sample and is shown as the dry curve on the graph. A third curve is also plotted, representing the half-value dry curve and is used to identify the average pore size, which occurs at the intersection between this curve and the wet curve. Until a pressure of 2 bar, the wet curve does not show any nitrogen flowrate; this is due to the gas pressure being unable to overcome the capillary pressure of the pores inside the sample, hence no liquid is pushed out. On the other hand, the dry curve shows a positive nitrogen flowrate at this range, which is expected since there is no wetting liquid present. The wet curve shows an initial steep increase at 2 bar pressure, as this is the first point where liquid is being pushed through the pores. Just before 2.5 bar, the wet and dry curves intersect, representing the pressure required to remove all the wetting liquid from the sample; therefore, this point identifies the smallest pore size. The smaller the pore

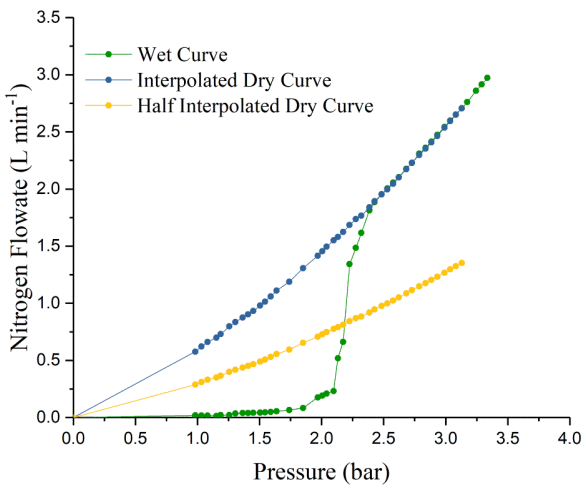


Figure 6: Nitrogen flowrate vs pressure (for freezing disc at 1150 °C)

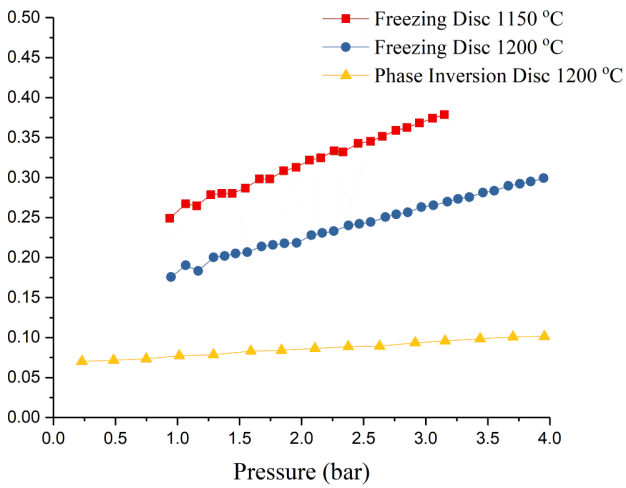


Figure 7: Nitrogen permeation flux vs pressure

size, the higher the hydrostatic pressure required to overcome the capillary pressure. The smallest pore size was chosen for comparison due to the large standard deviation of pore sizes obtained during the measurements and is shown in Table 1. The smallest pore size of the freezing membrane (1157 nm) is much bigger than that from the phase inversion membrane (794 nm). This meant that an average pore size would not be truly representative of the sample and hence comparisons would be less informative. The pore size distribution can be reduced by increasing the cooling rate, as this promotes homogeneous nucleation of the solvent, leading to a smaller pore size [17].

4.3. Nitrogen permeation flux

Figure 7 shows the relationship between nitrogen permeation flux and pressure for membrane discs obtained via phase inversion and freezing methods. A second membrane disc obtained via the freezing method was also investigated at a slightly lower sintering temperature. This graph uses data from the dry curve to obtain the nitrogen permeation flux. All the curves show a positive correlation between nitrogen permeation flux and pressure. However, the effect of pressure on freezing membranes is much more significant compared to phase inversion membranes. A pressure of 2 bar was taken for comparison between the curves. For the same sintering temperature, there is a significant increase in the nitrogen permeation flux when going from the phase inversion disc to the freezing disc, 0.09 to 0.22 mol m^{-2} s^{-1} Pa^{-1}. When comparing the two discs prepared via freezing at different sintering temperatures, it is clear that reducing the sintering temperature leads to a significant increase in the nitrogen permeation flux, 0.22 mol m^{-2} s^{-1} Pa^{-1} at 1200°C to 0.32 mol m^{-2} s^{-1} Pa^{-1} at 1150°C.

4.4. Water permeation flux

The results obtained from the phase inversion method and adapted freezing method are outlined in Table 1. The water permeation flux obtained from the freezing-membranes is almost double that from membranes produced via phase inversion. The increase in permeation flux can be explained by the difference between the inner structure of the phase inversion and adapted freezing membranes. The phase inversion discs form a very dense sponge like layer at both sides of the membrane surface which contributed to the majority of the water permeation resistance. Specific to phase inversion-NMP membranes, the middle micro-channels also showed a significant impact on the water permeation resistance [16]. Reflecting on the morphology of adapted freezing discs, the dendritic growth of the solvent crystals created non-uniform micro-channels across the whole membrane. The lack of a sponge like structure in the freezing ceramic discs leads to a significantly reduced water permeation resistance. This trend follows into the nitrogen permeation flux, with freezing showing an increase of more than two times when compared to phase inversion, as shown in Table 1. The dendritic structure produced through freezing can reduce the gas permeation resistance compared to phase inversion and the reduction in permeation resistance is more significant in gas than in water. This is due to the nature of water and nitrogen; water has a higher surface tension in the micro-channels, which requires a higher hydrostatic pressure to permeate through. On the other hand, nitrogen has a smaller particle size with a lower surface tension. The ceramic tip-structure of the freezing membranes offers a lower resistance to nitrogen particles.

Table 1: Numerical results obtained through experiments showing the comparison of properties between phase inversion and freezing membranes

Property	Phase Inversion-NMP	Freezing-DMSO
Water Permeation Flux (L m^{-2} hr^{-1} bar^{-1})	6,322	12,571
Nitrogen Permeation Flux (mol m^{-2} s^{-1} Pa^{-1})	0.085	0.222
Porosity (%)	54 ± 10	54 ± 10
Smallest Pore Size (nm)	794	1157

4.5 Porosity

The porosity calculated for both membranes was very similar. This is likely due to the limitations that were introduced through the method and equipment involved in determining this property. The porosity results are affected by the membrane thickness, membrane effective area and the wetting liquid used. Another factor that influenced this value was the difficulty faced in trying to completely remove the water from the surface of the disc without extracting water from inside the pore, creating a fluctuation for each result, even though repeat measurements were taken. Another technique to measure porosity involves mercury injection and would lead to more accurate results. This method involves immersing the disc in mercury which then enters inside the pores. The bulk volume can hence be determined by using the displacement of mercury. By increasing the pressure, the mercury is forced into all of the pores and the difference in mercury lost can be measured, allowing for the pore volume to be identified. The bulk and pore volume can be used to calculate the porosity of the disc. Due to resource limitations, it was not possible to measure the membrane porosity using this method [19].

6. Conclusions

An adapted method for the synthesis of porous ceramic membranes has been successfully implemented, utilising the foundation of the freeze-drying method and adapting it to improve energy efficiency. This was realised by replacing the primary drying step, previously involving sublimation under a vacuum, with an icy-water bath to remove the frozen solvent. The membranes produced via the freezing technique showed a significantly superior water and gas permeation flux compared to the phase inversion membranes at the same sintering conditions. The water permeation flux increased by 98% from phase inversion to freezing, whilst the nitrogen permeation flux increased by 161%. The freezing membranes also showed a much larger pore size with a very similar porosity.

The morphology of the membranes obtained through the adapted freezing method were analysed using SEM images and compared with those produced via the phase inversion method. The main difference was the lack of uniformity of the micro-channels present in the freezing membranes when compared to the horizontally uniform phase inversion membranes. The freezing membranes displayed a unique dendritic structure with ceramic tips. The substantial difference in the morphology of the two membranes opens up the potential for novel applications; one being catalytic membrane reactors as mentioned in (Gaudillere, C. and Serra, J. ,2016) [20].

7. Outlook

Time limitations were a major constraint when conducting these experiments. Repetitions of the experiments are required to improve the reliability of the results obtained. It is vital that the potential application of the membranes produced via the adapted freezing-technique are investigated further, such as coating the membrane with catalyst. The experiments performed were limited to a single freezing interface, stainless steel. This restricted our understanding of the influence of different freezing interfaces on the membrane morphology. A variety of freezing interfaces should be investigated to fully understand its effects. The composition of the suspension prepared for each experiment remained consistent. Hence, there is scope for varying the concentration of the solution and investigating the corresponding effects on the morphology and performance of the membranes produced. It was also identified that the sintering temperature can influence the nitrogen permeation flux. However, since the experiments were only performed at two different sintering temperatures, this evidence is not conclusive and requires further research.

Acknowledgements

We would like to extend our gratitude to Tao and Bo for sharing their expertise with the team and assisting in operating the PoroLux 1000 Porometer and SEM apparatus. We would also

like to thank Kang Li for offering us this opportunity to research under his support and guidance in a stimulating and inspiring environment.

References

[1] Grandviewresearch (2019). *Global Membrane Technology Market Size | Industry Report, 2019*. [Accessed 12 Dec. 2019].

[2] He X, Hägg MB. Membranes for environmentally friendly energy processes. *Membranes (Basel)*. 2012;2(4):706 – 726. Published 2012 Oct 18. doi:10.3390/membranes2040706 [Accessed 12 Dec. 2019].

[3] IETS. (2019). *Membranes as Energy-Efficient Technologies for Separation of Hydrocarbons | IETS*. [Accessed 12 Dec. 2019].

[4] Market, M. (2019). *Membrane Filtration Market by Application, Type, Region - Global Forecast 2025 | MarketsandMarkets*. [Accessed 12 Dec. 2019].

[5] Lee, M. (2016). *Micro-channel enhanced alumina membranes - designing and tailoring their properties for widened applications*. Ph.D. Imperial College of London. [Accessed 12 Dec. 2019].

[6] Synderfiltration. (2019). *Polymeric Membranes: Porous vs. Non-porous*. [Accessed 12 Dec. 2019].

[7] Bazzarelli, F., Giorno, L. and Piacentini, E. (2015). Porous Membranes. *Encyclopedia of Membranes*, pp.1-3. [Accessed 12 Dec. 2019].

[8] Baker, R. (2004). Membrane Technology and Applications. [Accessed 12 Dec. 2019].

[9] Ji, G. and Zhao, M. (2017). Membrane Separation Technology in Carbon Capture. *Recent Advances in Carbon Capture and Storage*. [Accessed 12 Dec. 2019].

[10] Ceramics.net. (2019). *Porous Ceramics*. [Accessed 12 Dec. 2019].

[11] Qian, L. and Zhang, H. (2010). Controlled freezing and freeze drying: a versatile route for porous and micro-/nano-structured materials. *Journal of Chemical Technology & Biotechnology*, 86(2), pp.172-184. [Accessed 12 Dec. 2019].

[12] Fukasawa, T., Ando, M., Ohji, T. and Kanzaki, S. (2001). Synthesis of Porous Ceramics with Complex Pore Structure by Freeze-Dry Processing. *Journal of the American Ceramic Society*, 84(1), pp.230-232. [Accessed 12 Dec. 2019].

[13] Araki, K. and Halloran, J. (2005). New Freeze-Casting Technique for Ceramics with Sublimable Vehicles. *Journal of the American Ceramic Society*, 87(10), pp.1859-1863. [Accessed 12 Dec. 2019].

[14] Wang, B., Ji, J. and Li, K. (2016). Crystal nuclei templated nanostructured membranes prepared by solvent crystallization and polymer migration. *Nature Communications*, 7(1). [Accessed 12 Dec. 2019].

[15] Liu, X., Wang, C., Wang, B. and Li, K. (2016). Novel Organic-Dehydration Membranes Prepared from Zirconium Metal-Organic Frameworks. *Advanced Functional Materials*, 27(3), p.1604311. [Accessed 12 Dec. 2019].

[16] Lee, M., Wang, B., Wu, Z. and Li, K. (2015). Formation of micro-channels in ceramic membranes – Spatial structure, simulation, and potential use in water treatment. *Journal of Membrane Science*, 483, pp.1-14. [Accessed 12 Dec. 2019].

[17] Wang, B., Ji, J., Chen, C. and Li, K. (2018). Porous membranes prepared by a combined crystallisation and diffusion (CCD) method: Study on formation mechanisms. *Journal of Membrane Science*, 548, pp.136-148. [Accessed 12 Dec. 2019].

[18] Creaparticlescience. (2019). *The POROLUX™ "true" First Bubble Point detection mode via Particle Science*. [Accessed 12 Dec. 2019].

[19] Wardlaw, N. and McKellar, M. (1981). Mercury porosimetry and the interpretation of pore geometry in sedimentary rocks and artificial models. *Powder Technology*, 29(1), pp.127-143. [Accessed 12 Dec. 2019].

[20] Gaudillere, C. and Serra, J. (2016). Freeze-casting: Fabrication of highly porous and hierarchical ceramic supports for energy applications. *Boletín de la Sociedad Española de Cerámica y Vidrio*, 55(2), pp.45-54. [Accessed 12 Dec. 2019].

An Environmental and Economic Assessment on the life cycle of Transport fuels: A Study on the Western Australian Heavy Good Vehicles Sector

Esther Adebajo, Yousuf Al-Riyami

Department of Chemical Engineering,Imperial College London, SW7 2AZ

Abstract

This paper investigates the possibility of low carbon transport fuels for heavy good vehicles (HGVs) in Australia. The study uses both a fuel life cycle assessment (LCA) and an economic assessment to find the fuel that can be substituted for the current market leader: diesel. The paper addresses the Western Australia's resources with fuel use modelled in the city of Perth. The economic assessment uses market prices for each fuel at the pump as a metric. To find a trade-off between environmental impact and economic feasibility, the Levelised cost of carbon mitigation (LCCM), a metric that factors the cost of reducing one Kg of CO_2 emissions with an alternative fuel to diesel, is used for each alternative fuel and benchmarked to diesel. Green gaseous hydrogen, produced from wind energy powered water electrolysis, has the lowest carbon dioxide (CO_2) life cycle emissions (LCE) at 0.008 Kg CO_2/MJ which accounted for a 90% reduction in emissions from diesel. Green Ammonia, produced from solar energy powered hydrogen production, was next with 0.035 Kg CO_2/MJ. LNG only resulted in a 31% reduction in of CO_2 emissions relative to diesel. On the price side, LNG has the lowest fuel price at \$0.016/MJ compared to \$0.023/MJ for diesel. The Green fuels both have an estimated price of \$0.056/MJ. The negative LCCM value for LNG, which accounts for a lower fuel price and LCE compared to LSD, makes it the fuel of choice for a short-term transition until green fuels become more price competitive. Methane emissions from LNG supply chains should be maintained below 1.5-3.5% of throughput for climate change benefits to be realised. It is also concluded that the vehicle fuel cell risk levels are similar to that of existing conventional fuels.

Keywords: Life Cycle Assessment, Heavy Good Vehicles, Western Australia Emissions, Greenhouse Gas Emissions, Liquefied Natural Gas, Ammonia, diesel, gaseous hydrogen

1. Introduction

1.1. Background

Global growth in GDP and commerce are leading a rise in demand for mobility and road freight. The freight transport sector is predicted to experience an increase in global transport volumes from 110 trillion tonne-kilometres in 2015 to 190 trillion tonne-kilometres by 2030 (1). Thus, the growing need for transport fuels around the world is contributing to an increase in oil demand, which is predicted to rise by 5 million barrels per day by 2050 for road freight transport. The environmental impacts of this are substantial, with a predicted increase in emissions of CO_2 by around 900 million tonnes through to 2050, the same emission growth as that of coal for use in both the power and industry sectors (2). Thus, technological intervention is necessary to curb the emissions arising from the road freight sector. Particularly with emissions from heavy good vehicles (HGVs), which are trucks with a gross vehicle weight greater than 3500 kg. HGVs represented 26% of all road vehicles in use in 2015 and contribute about 8% of global greenhouse gas emissions (3) (2).

Major economies such as the United States, China, Russia, EU, Japan, and Australia were responsible for emitting 68.5% of total global fossil CO_2 in 2018. From these countries, the largest CO_2 per capita emissions in 2018 (given in terms of tonne CO_2/cap) is found for Australia (16.77), followed by the United States (16.14), then Russia (12.14), Japan (9.42), China (7.95), UK (5.59), and India (1.94) (4). Despite Australia only contributing around 1.1% of global CO_2emissions, the high per capita emissions reveals that there is still room for improvement. Following the Paris Agreement, Australias climate target will be to reduce emissions to 26-28% on 2005 levels by 2030 (5). The transport sector is the second largest source of emissions in the Australian economy making up 18% of total emissions in 2017. Transport emissions are predicted to be 82% higher than 1990 level by 2030 – with emissions from HGVs expected to grow faster than any other transport emission (6). Thus, decarbonisation of the road freight sector is necessary for Australia to align with their Paris climate target.

1.2. Literature Review

Assessing previous work, arguments have been made for the feasibility of using low carbon fuels for HGVs through an environmental life cycle assessment (LCA). Many HGV studies have focused on Natural Gas (NG) in particular as a transportation fuel alternative (7; 8) in the UK and the US. Cooper et Al. (9), which studies a full suite of alternative fuels for comparison to NG, concludes that using natural gas as a fuel could help to cut greenhouse gas and air quality emissions in the sector but is not enough to meet the countrys 2030 climate targets. The paper also quantifies the environmental benefit of renewable fuels and finds it to be reduced for its high environmental impacts in land use change and other key categories that could negatively impact climate change efforts.

Beer et Al (10) centred a study on the Australian heavy vehicles sector, comparing the life cycle emissions of green-

Table 1: Summary of the case studies used for each fuel supply chain

Fuel	Extraction & Production	Distribution	Vehicle
Diesel	Cliffhead Facility (Crude oil Extraction) BP Kwinana refinery (Diesel Production)	Via trucks	Internal Combustion Engine (ICE)
LNG	Gorgon gas Plant (NG Sweetening) Onsite liquefaction (LNG production)	Via pipeline (through 9 compressor stations)	Spark Ignition engine
Ammonia	Dampier-Burnbury Pipeline (NG Supply) Yara Pilbara Plant (Ammonia Production)	Via Trucks	Fuel Cell Electric Vehicle (FCEV)
Gaseous Hydrogen	Port Lincoln Plant (Hydrogen Production)	Via Trucks (compressed for transport)	Fuel Cell Electric Vehicle (FCEV)

house gases from a range of alternative fuels which included Biodiesel, Liquefied Natural Gas (LNG) and Low Sulfur diesel. The LCA confirmed that renewable fuels (biodiesel and ethanol) can significantly reduce life cycle emissions. However, the findings from this study present an optimistic view with limited implementation considerations.

Addressing the economic impact of switching to low carbon fuels is important in the discussion on the real-world effects of this transition. Prevedouros et Al. (11) advocates for a more holistic view on low carbon transport but for seven light duty vehicles with different propulsion technology. By addressing the societal and consumer life cycle cost of each vehicle, the results showed a need to incentivise vehicles with low environmental impact (Fuel cell and Hybrid electric vehicles) through policies such as tax credits to increase their widespread adoption. Although these alternative fuel vehicles have now been commercialised and accepted for passenger vehicles, the same technology for HGVs is still in its early days.

This study combines an LCA with an economic assessment of diesel compared to alternative fuels for the HGV sector in Western Australia. This study considers only four fuels as the aim of this work is to weigh the financial penalty to the environmental impact of switching from diesel to some other fuels that have been proven ideal for HGV use as of present. The authors believe that this study will be of interest to transport policy makers and HGV manufacturers in Australia.

2. Methodology

2.1. Life Cycle Assessment

The LCA framework is used to determine the environmental impact of the HGVs powered by each of the four fuels considered. The LCA considers emissions across the fuel life cycle from Well-to-Wheel (WTW). The boundary for WTW begins from feedstock extraction and ends at vehicle use (see Figure 1). The greenhouse gas (GHGs) impact of carbon dioxide (CO2) emissions is the main criteria for comparison. The functional unit used is per MJ of engine uused by fuel as it is widely used in literature and allowed for comparisons. Four options are considered in total: diesel, Ammonia, LNG and Gaseous Hydrogen.

The justifications of fuel choice are as follows: firstly, Diesel represents 60% of fuel consumed by trucks in 2018 (12), and would therefore serve as a suitable standard and basis of comparison against the other fuels chosen. LNG is chosen due to Western Australia's large natural gas reserves, which have resulted in 66% of all Western Australian petroleum sales (13). LNG is the preferred fuel over compressed natural gas (CNG) for trucks, as LNG would occupy a much smaller volume in a given tank (1/600th the volume of natural gas) compared to CNG (1/100th only) thus reducing the size of vehicle on board storage tanks.

Green Ammonia and Hydrogen are chosen following incentives by the Australian government to create a hydrogen economy to address domestic needs for sustainable energy and transport (14). Hydrogen fuel is more suited to trucks than battery electric vehicle technology since hydrogen carries more energy than the equivalent weight of batteries.

2.2. LCA Modelling

A first principle approach to LCA is carried out in Excel. This is to gain full control of the input functions for the model in order to tailor it to the Australia context. The excel model also gives a clearer indication of the contribution at each stage of the life cycle. Emission factors from various literature sources are used to quantify the impacts at each stage of the supply chain relative to the functional unit of preference at each phase. Equation (1) shows an example of the method used to calculate the life cycle emissions on a "per MJ" basis for the refuelling station phase of the supply chain. Here $GHGLCE_{Fuelj}$ is the Greenhouse gas life cycle emissions "per MJ" of fuel used by engine for fuel j, EF is the Emissions Factor for the Australian electricity grid mix (assumed to be used to power the refuelling stations), TE is the annual electricity requirement for the refuelling stations in KWh and TM is the annual production of the fuel in MJ.

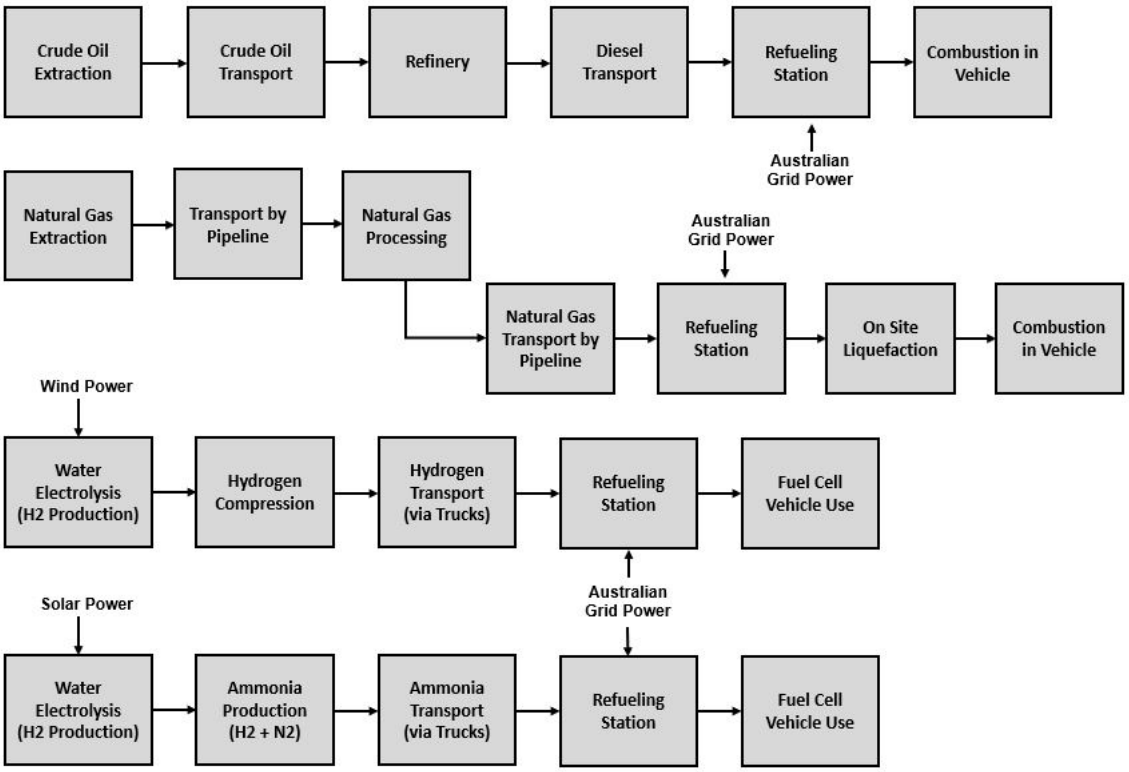


Figure 1: Flowchart of fuel life cycle of Diesel, LNG, Green Hydrogen and Green Ammonia

$$GHGLCE_{\text{Fuelj}}(kgCO_2/MJ) = \frac{EF(KgCO2/KWh) \times TE(KWh)}{TM(MJ_{\text{Fuelj}})} \quad (1)$$

The final Life cycle emission reported in this study is a summation of the per MJ CO_2 life cycle emissions contributions from each stage of the life cycle. The supply chain is also modelled on the Greenhouses Gases, Regulated Emissions and Energy use in Transportation (GREET) software for comparison purposes. The benefits of the two methods will be explored later on. Other background data sources that feed into the model are taken from the GREET database (15), Australian specific sources (if available) or estimated for the Australian context using international data (if unavailable). These are then modified for the Australia environment.

The HGVs are assumed to run 100% on the fuel used (no dual fuel engines considered). However, the impact of CH4 and NOx emissions are addressed for natural gas and diesel supply chain. The vehicle is limited to highway transport and, where possible, truck data to represent highway vehicles have been used. Environmental impact of the vehicle life cycle (vehicle manufacture to disposal) is not part of the scope of this study.

2.3. Levelized Cost of Carbon Mitigation (LCCM)

The LCCM metric, shown below in equation (2), is adapted from the 2018 study by Parkinson et Al. (16) which states- " The LCCM is an estimate of the additional cost associated with reducing the emissions of hydrogen production per tonne of CO2e against emissions of hydrogen production using [other technology]".

$$LCCM(\$/kgCO_2) = \frac{FinancialPenalty}{EnvironmentalGain} = \frac{Price_{\text{Fuelx}} - Price_{\text{LSD}}(\$/MJ)}{GHGLCE_{\text{LSD}} - GHGLCE_{\text{Fuelx}}(KgCO_2/MJ)} \quad (2)$$

In the original paper, the metric is used for comparing the relative costs of carbon mitigation for different hydrogen production techniques. In this paper, the metric is utilised to evaluate the relative economic and environmental merits of the fuels benchmarked to diesel (2). If a fuel is more expensive and its life cycle is more CO_2 intensive, then the LCCM value is not considered as this equates to a fuel that may not effectively be an alternative to diesel. LCCM is used in combination with other analysis to form the conclusions of this study.

2.4. Case Studies

The scope of the study is concentrated in Western Australia. All case studies see fuels being produced and supplied in WA where it is then distributed to and used in Perth, WA. To simplify the supply chains, case studies are picked for each fuel. These case studies are essentially single route suppliers and distributors in the whole fuel life cycle in this region. All refuelling station are assumed to be run on electricity from Australia's grid. Plants are expected to generate their own power on site with less grid input. The case studies are as follows

- **Diesel** – Crude oil is extracted from the Cliffhead facility. It is then transported via trucks to the BP Kwinana refin-

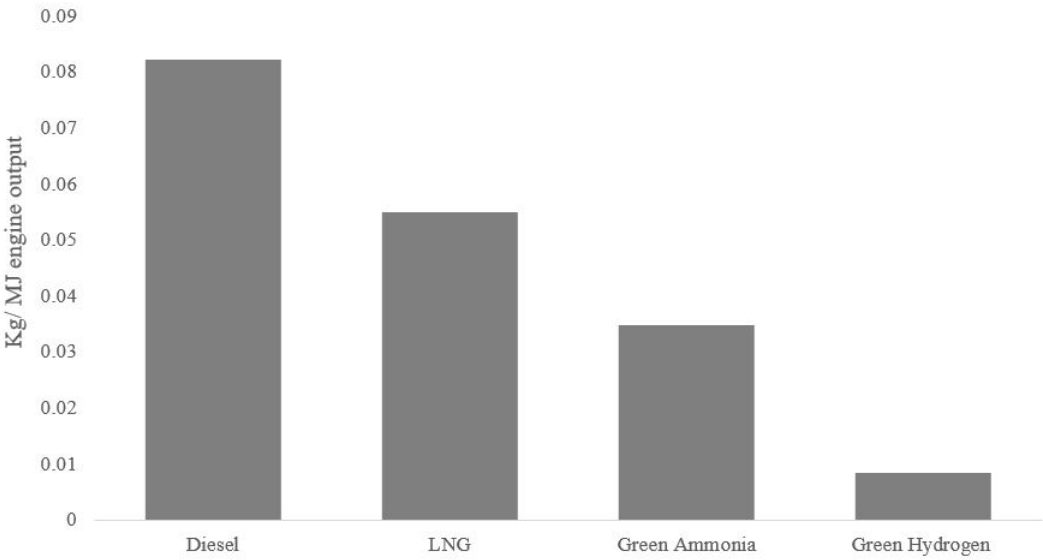


Figure 2: Life Cycle well-to-wheel CO_2 Emissions of all four fuels per MJ fuel used by engine

ery. From here, diesel is produced and transported to a Perth refuelling station. The HGV is assumed is assumed to have an internal combustion engine which produces GHGs tailpipe emissions.

- **LNG** — Natural gas is extracted from the Jansz-IO fields off the coast of western Australia. It is then transported via pipeline to the Chevron Gorgon Domestic gas plant (17) where the impurities, such as mercury, are stripped to sweeten the NG. The product is then transported via pipeline to the Dampier Burnbury pipeline (one of WA longest NG pipelines) which supplies Perth with NG. There are 9 compressor stations modelled along the pipeline between Gorgon gas plant input and Perth output. NG is then transported to the refuel station via trucks and is liquefied on site before being stored for distribution to vehicles. Vehicle is assumed to have a spark ignition engine which produces GHGs tailpipe emissions. Methane slip is also quantified for this supply chain route.

- **Renewable Ammonia** – NG is obtained from the Dampier pipeline. A connecting pipeline then sends the natural gas to the Yara Pilbara plant. This is one of the largest Ammonia producing plants in the word. A transition to renewable producing ammonia in this plant has been proposed ((18)) and so it is assumed the new renewable production facility is in operation for this study. Solar energy is used to produce electricity used by the electrolysers to produce hydrogen required as ammonia feedstock. After nitrogen then reacts with hydrogen to form ammonia, the product is sent via trucks to a refuel station in Perth. Ammonia is assumed to be a feasible fuel for fuel cell vehicles therefore there will be no GHG tailpipe emissions.

- **Renewable Gaseous Hydrogen** – Water is separated via wind powered electrolysis in the planned port Lincoln plant. This plant is not yet in operation therefore operations data is based on forecasts. This plant is also located in south Australia which has better wind energy potential than WA. The gaseous hydrogen product is compressed into cylinders and transported to a refuel station in Perth via trucks.

2.5. *Approach*

This study was carried out in five steps to achieve its aims

1. *Map out Fuel supply chain*; pick a supply chain case study in western Australia and ending in Perth
2. *A Life cycle inventory*; review current literature for energy, emissions and price specific data for each fuel supply chain
3. *Run LCA model and price analysis*; implement inventory data into Excel model and generate results for each fuel
4. *Use LCCM metric to evaluate technologies*; Identify the fuel with the highest potential based on LCCM results
5. *Sensitivity Analysis*; note the most impactful segments of the supply chain

3. Results and Discussion

3.1. *Well-to-Wheel CO_2 Emissions*

The results from Figure 2 show that diesel has the highest WTW followed by LNG, then green ammonia, and green hydrogen. This trend is expected, whereby the differences in WTW emissions of each fuel can be explained by the different technologies adopted by each fuel in their respective supply chains.

3.1.1. *Diesel vs LNG*

The emissions for diesel are found to be higher than that of LNG because of the higher tailpipe emissions of a diesel engine. The tailpipe emissions are 75 g CO_2/MJ and 55 g CO_2/MJ for diesel and LNG engines respectively (15; 19). The tailpipe emission figures are reasonable since the combustion emissions of natural gas are typically 75% that of diesel owing to diesels lower energy content and higher carbon content (20).

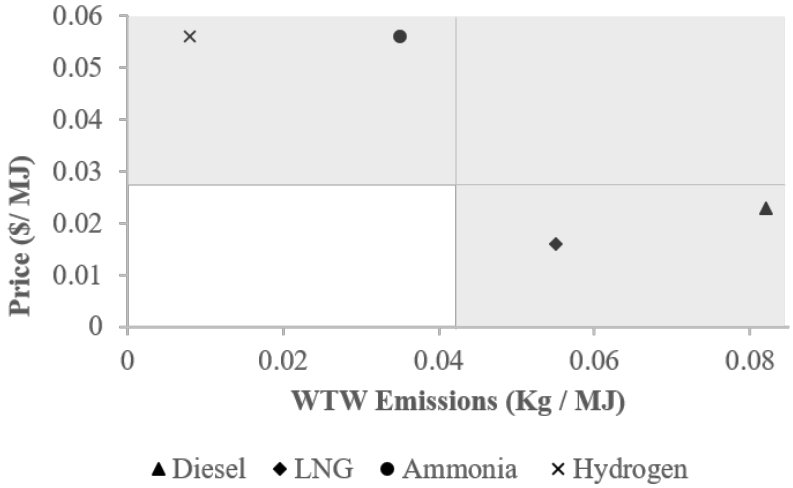


Figure 3: Economic and Environmental comparison of all four fuels highlighting the ideal fuel region in white.

3.1.2. LNG vs Green Ammonia

LNG has a higher WTW emission than that of green ammonia, since the direct CO_2 emissions from production of LNG are much greater than the indirect emissions associated with ammonia production through solar energy. The direct and indirect production emissions for LNG and ammonia facilities are 140 kt CO_2/ year and 21.8 kt CO_2/ year respectively (17) (21). The use of ammonia as a fuel is only justified through a renewable production route, as a conventional ammonia production route through steam methane reforming would result in production emissions being higher than that of LNG on a per product basis. They are are found to be 0.06 kg CO_2/ kg LNG from the Jansz I-O fields and 1.83 kg CO_2/ kg ammonia from the Steam methane reforming Yara Pilbara Plant (22) (23).

3.1.3. Green Ammonia vs Green Hydrogen

Green ammonia has a higher WTW emission than that of green hydrogen. There are two reasons for this: firstly, the production of green ammonia builds on initially producing green hydrogen from electrolysers, and then reacting that hydrogen with nitrogen in a reactor to produce ammonia. Thus a greater energy input per kg of product is involved in the process of green ammonia than producing green hydrogen. Secondly, the emission factor for solar energy is found to be 49 g CO_2/kwh (24) while that of wind energy is 25 g CO_2/kwh (25), this difference in emission factors can be explained by the fact that wind turbines are more efficient than solar panels and also come with a simpler structural makeup than the silicon and metals required to build a solar panel.

3.1.4. Per Kilometer Emissions

Out of the four fuels considered, only diesel and LNG contribute to emissions on the vehicle use stage. The fuel cell vehicles do not emit CO_2 into the atmosphere, thus an analysis on per vehicle emissions would not significantly alter the insights drawn from the per MJ functional unit. Moreover, the fuel consumption of diesel and LNG are very similar, averaging at 12.3 MJ/km and 14.4 MJ/km respectively (9), and would also vary as a function of the driving regime that the vehicle is subjected to (highway, urban, or rural driving).

Table 2: Comparison of investigation to literature

Fuel	Investigation (Kg CO_2/MJ)	Literature Results (Kg CO_2/MJ)
Diesel	0.082	0.085
Ammonia	0.035	0.040
Gaseous Hydrogen	0.008	0.040
LNG	0.055	0.069

3.2. Comparisons to Literature

Comparing the investigation results with literature, diesel is the fuel with an emissions value closest in line with its literature source. For LNG, the discrepancy in values arises from the fact that the investigation only considers presently available domestic resources that contribute to the life cycle emissions attributed to domestic LNG as a fuel. In contrast, the literature source considers the emissions associated with supplying natural gas (including extraction and processing) from a wide range of countries via imports as part of its domestic LNG well-to-wheel supply chain.

For hydrogen, although both the investigation and the literature source consider production of hydrogen using wind power, the main discrepancy arises from the phase of the hydrogen product. The literature source includes a liquefaction step to liquify gaseous hydrogen in production and transport liquid hydrogen to the refuelling station, while this investigation considers gaseous hydrogen without any liquefaction. Furthermore, an additional reason for the large difference between the emissions can be attributed to the fact that this investigation‘s supply chain considers that electricity for compression of hydrogen is obtained from on-site wind power in the facility as opposed to the grid in the literature source.

For Ammonia, since life cycle assessments have not yet been carried out considering a green ammonia route, the same literature source as that of green hydrogen is used as a basis of comparison. Nonetheless, the comparison is interesting to note as the life cycle emission values for producing ammonia and liquified hydrogen are very close. The higher life cycle emissions of liquid hydrogen than ammonia is due to the energy intensive nature of the liquefaction process, requiring a mean

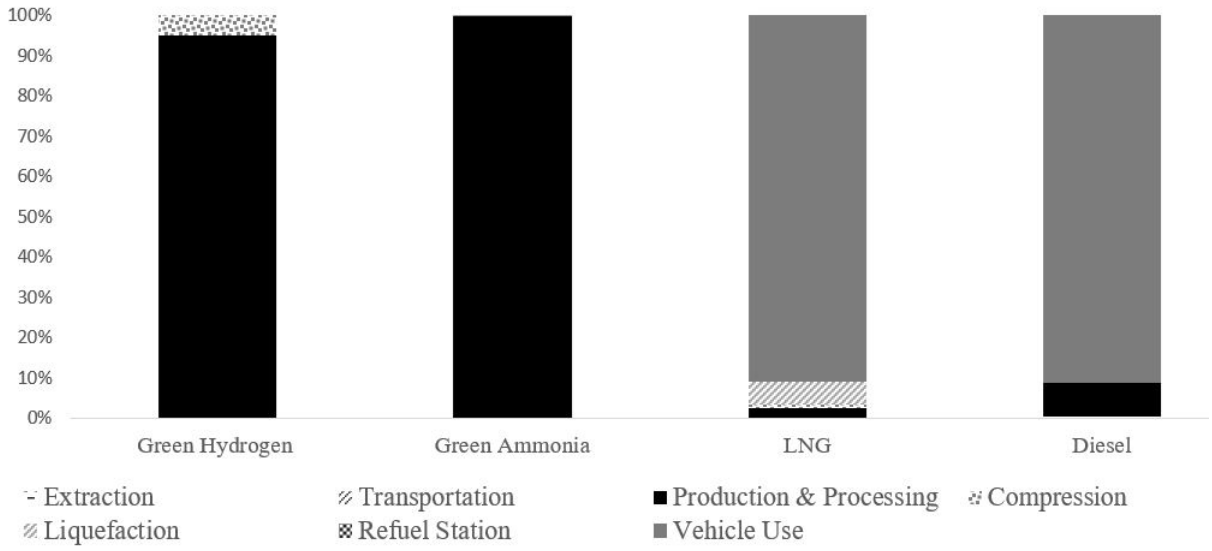


Figure 4: Percentage Supply Chain Contributions showing which stage of the fuel life cycle contributes what amount to the total emissions

electricity input of 10.9 Kwh/kg of H_2 liquified (26) compared to an ammonia electrolyser which uses a smaller energy input than that. For example, Haldor Topsoes solid oxide electrolyser requires 7.2 kwh/kg ammonia (27).

3.3. Greet Models

The supply chains of the four fuels are modelled on GREET. It should be noted that the production of ammonia on GREET was modelled using steam methane reforming from natural gas rather than through solar power. The WTW emission results obtained from GREET are summarised in Table 3. It is observed that the GREET results are not in line with literature results, particularly in the case of renewable hydrogen. The high emissions value for the hydrogen supply chain could be explained by the emission factor attributed to wind power being overestimated on GREET. The built-in data bank on GREET is useful for complex supply chains. Given that the supply chains considered are linear and do not contain overlapping pathways, a more robust first principle analysis is more suitable than the GREET software in this case.

3.4. Economic Assessment

3.4.1. Price-Emissions Comparisons

Fuel price is chosen as the metric to assess the economic considerations of the life cycle. The figure for price encompasses the scaling up of fuel costs throughout the supply chain up until the refuelling stage and is thus deemed appropriate.

In order to compare the fuel prices of diesel and LNG with the fuel prices of hydrogen and ammonia (which are not yet widely implemented) on a like-to-like basis for an Australian context, the fuel price is calculated in the following way:

$$\begin{aligned}&\text{Australian Fuel Price} =\\&\text{Price of fuel or commodity in the US} - \text{US Tax}+\\&\text{Fuel tax in Australia} + \text{Technology up-take price factor}\end{aligned} \tag{3}$$

Table 3: Comparison of GREET model to literature

Fuel	Greet Model (Kg CO_2/MJ)	Literature Results (Kg CO_2/MJ)
Diesel	0.075	0.085
Ammonia	0.20	0.069
Gaseous Hydrogen	0.13	0.040
LNG	0.052	0.069

The values for fuel price can be found in Table 4. It should be noted that since studies have not been conducted for pricing ammonia for transport fuel, the fuel price of green ammonia is assumed to be the same as the fuel price of hydrogen.

Plotting fuel price against emissions on a per MJ engine output basis, the graph is split into four quadrants. The quadrant split is carried out on the following basis: the horizontal line at 0.028 $\$/MJ$ represents a 50% reduction in the maximum price for renewable fuels and the horizontal line at 0.041 kg/MJ represents a 50% in the maximum emission seen for the non-renewable fuels. Thus, the bottom-left region in white in figure 3 represents the optimal fuel based on both environmental and economic considerations. A trade-off is observed as the four fuels are split in two distinct regions, with the green fuels coming at a high price but low emissions while the non-renewables are less expensive but have a greater environmental impact. Ideally, the desired is one that would be in the white region of Figure 3.

3.4.2. LCCM results

The results of the LCCM analysis show that there is a trade-off between the level of decarbonisation achieved and the cost of the fuel to the consumer. The higher the LCCM value, the higher the cost of carbon mitigation and thus the higher the price HGV users will pay at the pump. Therefore, LNG proves to be the most ideal fuel. LNG, although having a lower LCE than diesel, has a marginally lower per MJ price relative to diesel which results in a negative numerator of the LCCM formula and anoverall negatuve LCCM value. In this case, as the environmental impact is reduced, the LCCM result is favoured.

Table 4: LCCM and pump price results for all fuels

Fuel	Pump Price ($/MJ)	LCCM ($/Kg CO2)
Diesel	0.023	-
Ammonia	0.056	0.69
Gaseous Hydrogen	0.056	0.44
LNG	0.016	-0.26

From this analysis, it is clear that the environmental impact reduction attained by using renewable fuels as opposed to diesel is not large enough to offset the high cost of using this fuel. Given that such fuels require a high capital investment to set up the infrastructure needed for their production and distribution, the price at the pump is not yet competitive to that of diesel and LNG worldwide. Technological improvements have led to a fall in the cost of renewable energy according to the IREA (28). Therefore, continual investment in green infrastructure will ultimately decrease the cost of this energy and thus the cost of fuels made from this energy. Australia is not lacking in renewable energy sources. Western Australia has great solar potential whilst the south-west region has very consistent and high wind speeds- some of the highest in the world (29). Plans to grow the renewable infrastructure are starting to be put in place. One such example is the Yara Pilbara plant, a commercial plant planned to produce green ammonia from intermittent renewable resources. However, policy makers in Australia should seek to increase renewable energy investments for the transport sector to ensure that significant environmental benefits of switching to renewable fuels are not ceased by its cost.

3.5. Sensitivity Analysis

3.5.1. Supply Chain Contributions

Figure 4 shows the contributions of the different life cycle stages of each fuel. For both renewable fuels, hydrogen and Ammonia, the production and processing stage accounts for 97-99% of the total emissions.This is expected as there are no CO_2 emissions produced from using fuel cell engines. Production and processing for renewable fuels is not directly attributed to the reaction to make the product but the life cycle GHG emissions (LCE) for wind (15.3 gCO_2/kwh) and solar (66.4 gCO_2/kwh) power generation in Australia (26). These emissions account for the manufacturing and implementation of the wind and solar units as explained in section 3.1.3 (16).

Tailpipe Emissions from fuel combustion account for the largest contribution to LNG and diesel supply chains (90%). However, the actual GHG emissions from LNG is only 75% of that from diesel making it marginally better. Combustion of fossil fuels is a direct production route for GHG emissions which is expected to be most concentrated at the vehicle use. Similar to diesel trucks, LNG fuelled trucks can be fitted with a three-way catalyst in the exhaust to further reduce engine out emissions. The second largest impact for LNG is the liquefaction process (6%), production for diesel (8%) and compression for hydrogen (4%). These are all electricity intensive processes. Electricity can be taken from the grid or better yet generated on site (diesel) using natural gas. Power for the refueling station process is impacted by the countrys grid mix which, in Australia, contains a significant amount of coal (30).Therefore this will vary depending on the country studied. Gas compression electricity is assumed to be taken from wind power already generated for the electrolysis process at plant.

3.5.2. Price Sensitivity

The price of Australia‘s diesel is determined by the supply and demand balance in the Asia-Pacific market. Thus, the price of diesel is subject to fluctuation. With diesel prices on the rise; and the fact that around 75% of diesel is sold on a contract to commercial and industrial customers (31), there is incentive for customers to switch to using cheaper LNG fuel instead. Given tax fuel excise in Australia is the largest source of road-related revenue (32), an increase in the fuel excise of LNG from $0.287/kg to its diesel excise equivalent of $0.418/l can be expected as the uptake of LNG improves over time (33). This change would move LNG undesirably further away from the ideal fuel region in Figure 3.

If by 2030 an 80% adoption rate of hydrogen fuel cell electric vehicles occurs in Perth, as projected by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) (14), then the price gap in table 4 between diesel and hydrogen could disappear by 2037 as the hydrogen-powered vehicle market becomes more competitive and hydrogen extraction technologies improve (34). If the same trends are assumed for ammonia fuel cell vehicles, then both fuels may move in to the ideal fuel territory in Figure 3 within the next 20 years.

3.5.3. Global Warming Potential

Global Warming Potential (GWP) metric allows us to draw parallels between the different global warming impacts of different gases. It is benchmarked to the emissions of 1 ton of CO_2 and assesses how much energy 1 ton of gas will absorb over a 100-year (or 20 year) time period. Methane and NOx emissions can impact the global warming potential (GWP) of the different fuel supply chains in this study. Although the focus of this paper is to assess the CO_2 life cycle emissions, the impact of other key GHG emissions will be discussed in this section as their GWP can lead to a more potent warming of the earth than CO_2 if not managed well. According to the EPA (35), 1 ton of Methane and NOx release is equivalent to the release of 28-36 and 265-298 tons of CO_2 respectively. In addition, methane does not persist in the atmosphere for as long as CO_2 or NO_X, so the strength of methanes warming is reduced over time (36). Overall, as Cooper et Al.(9) concluded, engines need to be better maintained to improve the onboard fuel systems efficiency and to reduce tailpipe emissions.

3.5.4. Methane Emissions

Methane emissions have been well documented in literature for natural gas fuel supply chains. Balcombe et Al (25) noted that methane can be released from different sources across the supply chain: leakages in pipelines, on-board fuel storage and vehicle tailpipe and crankcase emissions. Cooper et Al (9) quantified the methane emissions along the supply chain as a percentage of throughput at each stage for both LNG import and onsite liquefaction routes. The fuel station emissions are recorded as the lowest percentage (0.00007%) of throughput of each stage. These low emissions can be a result of a quick throughput through stations leading to little or no time for LNG boil-off to be initiated and also good maintenance of machines

and systems at refuelling stations. Truck methane emissions are the highest at an average of 0.55% of throughput at each stage. Vehicle fuel tank venting is considered as the intensity of HGV driving leads to full use of fuel within the 5-day minimum holding time before boil-off is intitated (37). Therefore, Methane emissions at this stage come from unburnt methane passing through the tailpipe or escaping through the engine crankcase during combustion. Slip emissions are recorded to range from as low as 0.2 g CH_4/km to as high as 1.65 g CH_4/Km and are dependent on factors such as vehicle fuel consumption, age of vehicle and engine temperature (38). To reduce this impact, catalysts can be used to limit the tailpipe methane emissions. Examples include a three-way catalysts or oxidation catalysts. Balcombe et Al. (25) explores vehicle methane emissions in more detail. Overall, the impact of methane emissions has been found to be predominately from trucks. Although the value is still much lower than the 3.00-3.5% (9) of throughput maximum to ensure a climate benefit of NG fuels compared to diesel. It is important to note that all values stated in this chapter are for a per km functional unit.

3.5.5. N_2O and NO_X Emissions

Balcombe et Al (25) quantified the effects of NO_X emissions from natural gas and diesel HGVs. N_2O has the highest GWP of methane and diesel which in turn can cause stronger global warming effects. These emissions are dependent on a variety of factors which include combustion temperatures and emissions control systems. Given the complexity of the reactions, the data on N_2O emission for natural gas vehicles is limited by lack of research. Therefore this study cannot account for the impact of N_2O emissions with certainty. A European study in the Netherlands carried out a tank-to-wheel Emissions testing of two Euro VI LNG heavy-duty vehicles against that of a diesel truck (39). The findings from this study is used to quantify the diesel and LNG truck NO_X emissions by other literature sources of the kind including Balcombe ET AL (25). The study found that NOx emissions from both LNG trucks, for an average vehicle performance (accounting for urban/rural/highway driving), is equal to or greater than the average of the 6 tested diesel trucks. Specifically, highway driving, which is assumed to be the predominant driving mode in this paper, has the lowest emissions for the diesel and LNG trucks tested. Cold engine starts are also found to account for a higher proportion ofNO_X emissions from diesel trucks than LNG trucks as the selective catalytic reduction (SCR) needs time to warm up to reach optimal operating conditions. In addition to this, the performance of the SCR is noted by Cai et al (40) to significantly increase the tailpipe emissions from diesel vehicles as opposed to NG vehicles.

3.6. Safety Considerations

Results from safety assessments conducted on the uptake of ammonia as a fuel reveal that ammonia has the lowest probability of an accidental release in truck transport (41). This can be ensured with the refrigeration of ammonia in pressure tanks during road transport to refuelling stations. Moreover, leaks and ruptures of hoses at the refuelling station can cause serious danger through small but long lasting releases. Thus, a safety distance of 150 m from public areas should be maintained to reduce public leakage impacts of ammonia (41). When it comes to the potential for collision accidents, the safety assessments reveal that the fatality rate of motor-vehicle accidents is not influenced by the fuel used to power the vehicle (42). Therefore, following the same reasoning for adopting renewable hydrogen, both fuels are deemed appropriate for uptake when suitable engineering precaution measures are taken across their supply chains.

4. Conclusion

4.1. Conclusion

This study carried out a life cycle assessment in combination with an economic assessment to compare diesel to three alternative fuels in the setting of the Australian market. The aim of this report is to realistically assess how the heavy good vehicles sector can reduce its emissions which is currently the second largest source of emissions in the Australian economy. The three fuels benchmarked to diesel are LNG, green ammonia and green gaseous hydrogen. The LCA and price results are classified on a per MJ engine output basis and are limited to this functional unit in this study. The effect of fuel price on fuel adoption rates for the consumer is assessed. The LCA of the three alternative fuels all had lower LCE than diesel showing a positive environmental benefit to switching, even when considered on a life cycle basis. The LCA favoured renewable fuels with the wind energy produced gaseous hydrogen having the lowest total CO_2 emissions along the supply chain. This is attributed to the lower emission factor for wind energy compared to that of solar energy and the less energy intensive production process for hydrogen. Green ammonia is the next best fuel because of its renewable production route which uses solar energy. LNG emissions make up 75% of diesels total LCE making it the third best fuel. The difference in tailpipe emissions between diesel and LNG explains this result. The economic assessment is measured using fuel price at the pump. Based on todays prices, LNG has the lowest pump price followed by diesel. Ammonia and hydrogen prices are nearly three times as high as diesel prices as these fuels are yet to be widely commercialised in Australia and are expensive to produce given the lack of infrastructure. Two metrics are used to quantify the optimal fuel based on a trade-off between the environmental and economic impact of switching fuels: LCCM and a fuel vs emissions plot. Both metrics are skewed towards LNG for having lower emissions and a lower price than diesel. The LCCM value itself is negative which indicates no financial penalty with LNG use.

The sensitivity analysis shows that tailpipe emissions accounted for an average of 94% of the supply chain contributions for LNG and Diesel individually. For renewable fuels, production and processing accounted for more than 90% of supply chain contributions. Contributions from other GHG emissions along the diesel and LNG supply chains are also accounted for in the sensitivity analysis. Methane emissions from the LNG route are recorded to be lower than the limit of 1.4-3.5% of throughput required for LNG to not be disqualified by methanes GWP. This suggests that the overall global warming impact of methane slip from truck engines, which makes up a large proportion of supply chain methane emissions, is not significant enough to change the results of this study.NO_2, which has a

very high GWP, and NOx emissions can be controlled by fitting truck engines with catalytic systems that reduce their emissions. However, catalytic systems and truck engines need to be maintained for their effective performance. Safety considerations for ammonia and hydrogen as fuels reveal that their risk levels are acceptable and similar to current fuels, and thus should not hinder fuel adoption if appropriate safety precautions are also implemented.

4.2. Outlook

Overall, this LCA successfully identified LNG as a readily available present-day solution to reduce Western Australias HGV sector emissions with no financial penalty. However, given LNGs marginal emission reductions, it should be seen as a short-term viable solution until Australia can transition to the technologies of the future which use renewably produced fuels in fuel cell vehicles. This conclusion is also expressed by Cooper et Al (9) for the UK context.

4.3. Further Work and Limitations

First, the choice of case studies focused on WA gives only a minute perspective of the potential for wider Australia. Further research can look into how the results differ by regions in Australia. The impact of more supply chain inputs at each stage of the life cycle will also further support the conclusions from this study. For example, including the impact of feed-stock imports for LNG and diesel supply chains.

Second, a larger life cycle inventory can be utilised to accurately evaluate an error range for the life cycle emissions, price and more literature comparisons. However, the small pool of Australia specific literature and information is a limiting factor for this study. As the gaseous hydrogen production plant is not yet in operation, the emissions associated with this route is based on predictions and estimations based on worldwide plants of a similar size. Due to the sparse data on ammonia as a transportation fuel in literature, it is difficult to do a like-for-like comparison for this supply chain with literature or data from other parts of the world. The safety considerations and competition with the lucrative fertiliser industry for ammonia also presents a hurdle to ammonia adoption as a transportation fuel. More work should be done to test the feasibility of this fuel for HGVs. This study also uses hydrogen fuel prices to predict ammonia fuel prices. However, ammonias prices could be lower given the already available sector.

References

[1] U. S. Development, "Global mobility report 2017," *Open Knowledge Repository*, Oct 2017.

[2] J. Teter, P. Cazzola, T. Gul, E. Mulholland, P. Le Feuvre, S. Bennett, P. Hugues, Z. Lagarde, V. Kraayvanger, T. Bryant, *et al.*, "The future of trucks: implications for energy and the environment," 2017.

[3] I. Wagner, "Number of vehicles in use worldwide 2015," Nov 2018.

[4] M. Crippa, G. Oreggioni, D. Guizzardi, M. Muntean, E. Schaaf, E. Lo Vullo, E. Solazzo, F. Monforti-Ferrario, J. Olivier, and E. Vignati, "Fossil co2 and ghg emissions of all world countries," 2019.

[5] "Australia's 2030 climate change target," 2015.

[6] "Australia's vehicle fleet - dirty and falling further behind," Mar 2019.

[7] H. Cai, A. Burnham, R. Chen, and M. Wang, "Wells to wheels: Environmental implications of natural gas as a transportation fuel," *Energy Policy*, vol. 109, pp. 565–578, 2017.

[8] A. Arteconi, C. Brandoni, D. Evangelista, and F. Polonara, "Life-cycle greenhouse gas analysis of lng as a heavy vehicle fuel in europe," *Applied Energy*, vol. 87, no. 6, pp. 2005–2013, 2010.

[9] J. Cooper, A. Hawkes, and P. Balcombe, "Life cycle environmental impacts of natural gas drivetrains used in uk road freighting and impacts to uk emission targets," *Science of The Total Environment*, vol. 674, pp. 482–493, 2019.

[10] T. Beer, T. Grant, D. Williams, and H. Watson, "Fuel-cycle greenhouse gas emissions from alternative fuels in australian heavy vehicles," *Atmospheric environment*, vol. 36, no. 4, pp. 753–763, 2002.

[11] P. Prevedouros and L. Mitropoulos, "Life cycle emissions and cost study of light duty vehicles," *Transportation Research Procedia*, vol. 15, pp. 749–760, 2016.

[12]

[13] "Natural gas - lng," Apr 2018.

[14] S. Bruce, M. Temminghoff, J. Hayward, E. Schmidt, C. Munnings, D. Palfreyman, and P. Hartley, "National hydrogen roadmap," 2018.

[15] "Argonne national laboratory," 2017.

[16] B. Parkinson, P. Balcombe, J. Speirs, A. Hawkes, and K. Hellgardt, "Levelized cost of co 2 mitigation from hydrogen production routes," *Energy & Environmental Science*, vol. 12, no. 1, pp. 19–40, 2019.

[17] "Gorgon gas development and jansz feed gas pipeline," May 2015.

[18]

[19] R. Wurster, W. Weindorf, W. Zittel, P. Schmidt, C. Heidt, U. Lambrecht, A. Lischke, and S. Müller, "Lng as an alternative fuel for the operation of ships and heavy-duty vehicles," 2014.

[20] "U.s. energy information administration - eia - independent statistics and analysis," Jun 2019.

[21] T. W. Australian, "Yara aims to move wa into green hydrogen," Feb 2019.

[22] "National inventory submissions 2019," May 2019.

[23] "Ammonia plant, burrup peninsula - epa wa," Aug 2015.

[24] D. Nugent and B. K. Sovacool, "Assessing the lifecycle greenhouse gas emissions from solar pv and wind energy: A critical meta-survey," *Energy Policy*, vol. 65, pp. 229–244, 2014.

[25] P. Balcombe, J. Speirs, E. Johnson, J. Martin, N. Brandon, and A. Hawkes, "The carbon credentials of hydrogen gas networks and supply chains," *Renewable and Sustainable Energy Reviews*, vol. 91, pp. 1077–1088, 2018.

[26] A. Ozawa, M. Inoue, N. Kitagawa, R. Muramatsu, Y. Anzai, Y. Genchi, and Y. Kudoh, "Assessing uncertainties of well-to-tank greenhouse gas emissions from hydrogen supply chains," *Sustainability*, vol. 9, no. 7, p. 1101, 2017.

[27] T. Brown, "Green ammonia: Haldor topsoe's solid oxide electrolyzer," Mar 2019.

[28]

[29] P. Newman, A. Went, and J. Wal, "Renewable transport: How renewable energy and electric vehicles using vehicle to grid technology can make carbon free urban development.," 2008.

[30] 2019.

[31] "Facts about diesel prices and the australian fuel market," Dec 2019.

[32] C. Parliament and P. House, "Revenue from road use," May 2017.

[33] A. T. Office, "Excise rates for fuel," Jul 2019.

[34] J. Arnott, "What would it cost to run a city on hydrogen?," Oct 2019.

[35]

[36] D. M. Cain, "A new way to assess global warming potential of short-lived pollutants," 2019.

[37] L. Gunnarsson and E. Helander, "How to handle boil-off gases from lng trucks," 2015.

[38] J. Norris, C. Brannigan, B. White, A. Demurtas, A. Sale, and I. Dawson, "Provision of hgv emissions testing," 2015.

[39] R. Vermeulen, R. Verbeek, S. van Goethem, and R. Smokers, *Emissions testing of two Euro VI LNG heavy-duty vehicles in the Netherlands: tank-to-wheel emissions*. Den Haag: TNO, 2017.

[40] H. Cai, A. Burnham, R. Chen, and M. Wang, "Wells to wheels: Environmental implications of natural gas as a transportation fuel," *Energy Policy*, vol. 109, pp. 565–578, 2017.

[41] N. J. Duijm, F. Markert, and J. L. Paulsen, "Safety assessment of ammonia as a transport fuel," 2005.

[42] "Comparative quantitative risk analysis of motor gasoline ...," Jun 2009.

Modelling oscillatory flow in helical reactors using adaptive mesh refinement simulations at low Reynolds numbers

Abrar Masud and Muhammad Waabis

Department of Chemical Engineering, Imperial College London, UK

Abstract

Sophisticated flow problems can be solved via the Basilisk solver by using a novel feature called adaptive mesh refinement (AMR), where complex geometries can be embedded on Cartesian meshes with implicit shape equations. In this study, we report the use of AMR to simulate the conditions of oscillatory flow experiments conducted in helical tubes. A qualitative agreement was found for the influence of oscillation frequency and amplitude on plug-flow performance in a two-dimensional reactor geometry. Varying both parameters gave a 'switch-on' and 'switch-off' effect, as observed previously in the experimental literature. The optimum conditions; which produced the highest tanks-in-series parameter, N; were found to be at 3 Hz frequency and 2 mm amplitude. Experimental results at lower Reynolds number, Re, were also validated and exponentially increased plug-flow performance was observed. Conditions of $\mathrm{Re} = 5$ gave a value of $N = 59$, whereas conditions of $\mathrm{Re} = 100$ gave a value of $N = 27$. Although a qualitative agreement was found, the findings cannot be directly compared to experiments as the simulations were two-dimensional. However, the results from simulations suggest there is an analogy to the Dean vortices, which can be scaled down to two-dimensional flows. Three-dimensional geometries were developed as a proof of concept, but oscillatory flows were not tested due to time and computational constraints.

1. Introduction

1.1 Plug flow

Plug flow is a desirable condition in tubular reactors, characterised by strong radial mixing across the channel and negligible axial mixing. It allows for better control of process variables and reduces operating costs. However, generally there is backmixing and axial dispersion which interferes with ideal plug flow conditions. Axial dispersion leads to non-idealities, making the reactor system less efficient than ideal plug flow (Reay et al., 2013; Levenspiel, 1999).

Various techniques have been explored to improve mixing in tubular reactors. Commonly used techniques are oscillatory baffled reactors (OBR) and coiled flow inverters (CFI). OBRs consist of a tube, typically in a U-tube shape, fitted with baffles that are equally spaced across the channel (Harvey et al., 2019). The baffles are fitted transversely to an oscillatory flow. This technique has seen various applications in heat exchangers (Law et al., 2018), biofuels production (Masngut et al., 2010), polymerisations (Parida et al., 2014), and wastewater treatment – all applications where strict control of flow is important. OBRs have also been developed in mesoscale sizes, with mL/hr throughputs (characteristic diameter of 4.4–5 mm) for process development applications. CFIs are tubular reactors as helical coils, often with periodic bends. Bends change the direction of the centrifugal force, giving rise to Dean vortices that encourage radial mixing. Applications of CFIs include protein refolding (Sharma et al., 2016), polymerisation, heat exchangers, and industrial mixers.

1.2 Dean vortices

In order to understand why oscillating fluid behaviour in helical coils can achieve ideal PFR conditions, we must first study the formation of Dean vortices (Ligrani, 1994; Nivedita et al., 2017) in coiled tubes. Instead of having a laminar radial velocity profile, liquids flowing in curved tubes experience centrifugal forces and the maximum radial velocity shifts towards the outer side of the curve. This centrifugal force is balanced by a pressure gradient which increases the pressure at the outer wall (further from the centre of curvature) relative to the inner wall (closer to the centre of curvature). This is referred to as the Dean instability (McDonough et al., 2019). Due to the no-slip boundary condition, the liquid at the top and bottom walls has a lower velocity than the liquid at the centre of the tube. This implies that there is a lower centrifugal force near the walls of the tube relative to the centre, hence requiring a smaller pressure gradient to balance this outward motion. Beyond a critical liquid velocity, the pressure gradient will dominate the centrifugal force at the tube wall, resulting in a liquid flow from the top and bottom

walls towards the centre of the tube. This flow is balanced by a restoring flow from the centre of the tube to the outer walls. This secondary flow develops as a pair of counter-rotating vortices, known as Dean vortices. The circulation magnitudes of the vortices increase as the curvature of the tube increases.

A dimensionless Dean number can be used to characterise these Dean vortices. The formula for Dean number as shown in eqn (1) accounts for fluid inertia, centrifugal forces and internal viscous dissipation. Dual vortex pairs are typically observed for $\mathrm{De} > 64-75$ while additional secondary instabilities in the form of vortex pair undulations, vortex pair twisting, and vortex splitting/merging can be observed for $\mathrm{De} = 75-200$. The Dean number is given by

$$\mathrm{De} = \sqrt{\frac{\frac{1}{2}\times\text{Inertial forces}\times\text{Centrifugal forces}}{\text{Viscous forces}}} = \mathrm{Re}\sqrt{\frac{D_t}{2R_c}} \quad (1)$$

where $$\mathrm{Re} = \frac{\rho v\, D_t}{\mu} \quad (2)$$

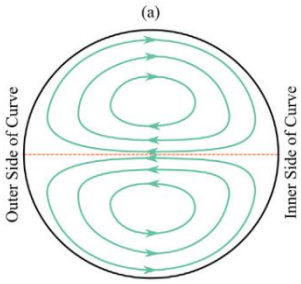


Figure 1. Dean Vortices visualised in a pipe (McDonough et al., 2019)

1.3 McDonough (2019) experiment

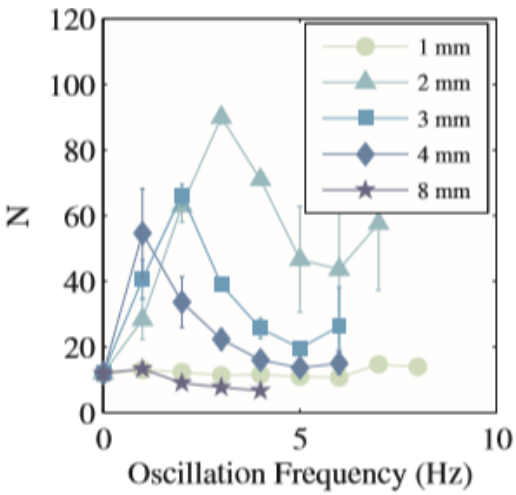


Figure 2. Tank-in-series model, N, versus oscillation frequency for various amplitudes (McDonough et al., 2019)

McDonough (2019) have combined the main elements of CFIs and OBRs, consisting of a tubular system with oscillatory flow inside a helical coil. Experimental results suggest that the setup achieved high degrees of plug flow in laminar conditions. The special helical reactors were manufactured with 3D-printing techniques, where radius of curvature, tube pitch and number of turns were varied. Plug flow conditions were tested by injecting a step tracer into the system and evaluating its residence time distribution (RTD). For ideal or near-ideal plug flow, the distribution for a fluid element will have a narrow peak and be symmetrical. In contrast, a wide distribution suggests deviation from ideal plug flow behaviour, hence significant axial dispersion. A tanks-in-series model was applied to quantify reactor performance. The parameter N, which denotes equivalent tanks-in-series, also determines the level of plug flow. Tanks-in-series relates to RTD as such (Fogler, 2016):

$$E(\theta) = \frac{N(N\theta)^{N-1}}{(N-1)!}e^{-N\theta} \quad (3)$$

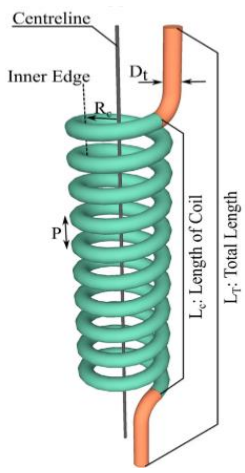


Figure 3. 3D model of the helical coil printed by McDonough, 2019

Oscillation frequency and amplitude of the flow were varied, and the corresponding N values were computed. For almost all conditions, the minimum value for equivalent tanks-in-series was found to be 10, indicating an acceptable level of plug flow. Values of $N > 100$ were also observed for some operating conditions, suggesting a good enhancement of plug flow behaviour. This is a significant improvement of the meso-OBR, where maximum N values reported in literature ranged from 25-50 depending on baffle geometry. Increasing oscillation frequency has 'switch-on' and

'switch-off' characteristics, showing a clear optimum value of N. A typical profile for a set Reynolds number ($\mathrm{Re} = 10$) is given by figure 2 and figure 4.

One explanation for this phenomenon was due to the fact that increasing oscillations gave rise to Dean vortices, which encouraged radial mixing. This corresponded to 'Oscillatory Dean numbers' of 70–200. However, Dean numbers above this region lead to a decrease in plug flow performance. Furthermore, it was also found that an oscillation amplitude of 2 mm, and a frequency 3 Hz gave a global maximum from all conditions tested, regardless of Reynolds number. This resulted in equivalent tanks-in-series of 100 (McDonough et al., 2019). Further parameters such as tube pitch, radius of curvature and flow rate were also tested, however this is outside the scope of the present project.

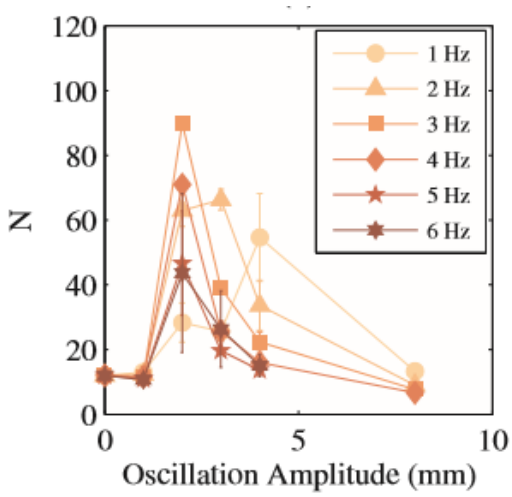


Figure 4. Tank-in-series model, N, vs Oscillation amplitude for various frequencies, (McDonough et al., 2019)

1.4 Basilisk

The scope of this project is to simulate this behaviour, as a computational fluid dynamics problem, and replicate the results with the intention that the simulation setup will be exploited for shape and parameter optimisation studies in the future. Basilisk (Popinet, 2016; Popinet, 2019) is an open-source code that solves partial differential equations on adaptive cartesian meshes. Therefore, it is quite useful to model complex mathematically defined geometries such as helices and integrate machine-learning and optimisation algorithms to the results of the simulations to determine the optimal input parameters. Further benefits, and some drawbacks, of using Basilisk will also be discussed. Basilisk has been successful in modelling many applications such as porous media and vortex-shedding problems, and these existing models will be built upon. The ultimate objective is to reproduce the data from McDonough et al. (2019) through modelling. Moreover, it is also aimed to simulate conditions that were deemed impractical for experimentations, leading to geometry optimisation studies.

2. Methodology

2.1 Adaptive mesh refinement

Two-dimensional and three-dimensional meshes were developed using the Basilisk (Popinet, 2019) solver. The main benefit of using this software is that it allows the use of AMR (Van Hooft et al., 2018). This novel feature is a significant advantage over static meshes to ensure computations can be done over various length scales. AMR means that cells can be finer in targeted regions, e.g. close to the walls, and coarser in less dynamic regions. The mesh is adapted with the solution at each timestep, and generally results in faster computations. Basilisk uses quadtrees in two dimensions for efficient grid refinement. Each node has exactly four children, whose partition can be triggered according to discretisation error. For three dimensions, oct-tree grids are used, where each node has eight children (Fuster et al., 2009). Figure 5 and 6 show examples of adaptive mesh refinement in both the construction of geometry, and in a velocity field. It is important to note that refining meshes with each timestep of the computed solution is particularly useful for the modelling of complex flows and vortices.

To achieve this, the following code was used:

```
event adapt (i++) {
  adapt_wavelet ({h}, (double []){4e-3},
maxlevel = 8);
```

This instructs Basilisk to refine the resolution according to the discretisation error defined by `h`. The field `h` in this model are the volume fraction for fluid volume, velocity and tracer concentration. When the discretisation error is larger than 4×10^{-3}, the mesh is refined by partitioning the parent cells into children cells. This is very useful for walls, phase boundaries, and vortices. The minimum and maximum quadtree level of refinement can also be set, defining the number of partitions Basilisk can make. Larger numbers of partition can potentially make the solution to complex problems converge but can be computationally expensive. The values of discretisation errors and levels of refinement for each variable are provided in Supplementary Information in Section 6, alongside the full Basilisk code. It is important to note that lower values of discretisation error and refinement were used in 3D

simulations in order to manage the heavier computational requirement.

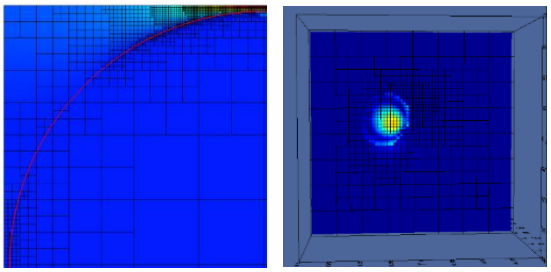


Figure 5 (left). AMR around a defined interface

Figure 6 (right). AMR around a flow field

2.2 Simulating fluid flow: Defining fluid in cells

Basilisk employs the finite-volume method to solve the Navier–Stokes (NS) equation. A function ϕ (`phi`) is defined which approximates the surfaces between two phases. Basilisk has general multiphase flow applications, however in the present scenario the two phases are fluid and wall (no fluid). The function ϕ is taken as input, and converted into a variable `cs`. This is done through interpolation between adjacent phi values. The variable `cs` denotes whether the cell is active (i.e. whether fluid is present), with its value going to zero at the walls. For coding purposes, positive values of `phi` at set co-ordinates imply that the NS equation is solved in the cell. Whereas negative values of `phi` mean no flow is present due to a wall.

2.3 Simulating fluid flow: Channels, coils and helical tubes

A scale-up approach was taken to create the simulations. Starting from basic flows, with code available from open-source examples (Thoraval et al., 2013), the complexity and realism were gradually increased until a three-dimensional helix was achieved. The first iteration of model development was a two-dimensional channel. This was defined by an intersection of two lines, such that the function ϕ would be positive for $-0.5 \leq y \leq 0.5$:

```
phi[] = intersection (0.5 - y, 0.5 + y) ;
```

The next step in complexity was a two-dimensional coil, which was a 2D projection of a helical tube. A similar technique was employed as for the channel, where the function ϕ was replaced with the intersection of two sinusoidal functions:

```
phi[] = intersection (0.25+0.25*sin(5*x) -
y, y + (0.25-0.25*sin(5*x))) ;
```

This gives a pipe thickness of 0.5 units, and since 1 unit is scaled to be 1 cm, this matches McDonough (2019) experimental condition of 5 mm tubes. The tube pitch was also scaled with the same unit length in the domain.

The final increase in complexity was to develop a helical structure in a three-dimensional mesh. Several methods were attempted, including parametric mapping, and importing CAD files. However, the simplest and easiest method to implement was to define the geometry implicitly. Equation (4) gives the implicit equation of a helix, such that the function `phi` will positive inside the tubes, and negative outside.

$$\phi = r^2 - \left(y - R\cos\left(\frac{x}{a}\right)\right)^2 - \left(z - R\sin\left(\frac{x}{a}\right)\right)^2 \tag{4}$$

Note that x, y, and z are Cartesian co-ordinates and the equation outputs a helix in the x-direction. There are three characteristic parameters. R is the distance from the centre of the helix to inner edge, r is the inner radius of coil tubes, and a is a parameter proportional to the tube pitch. Lower values of a give more turns in the coil.

2.4 Boundary conditions: Constant velocity Poiseuille flow

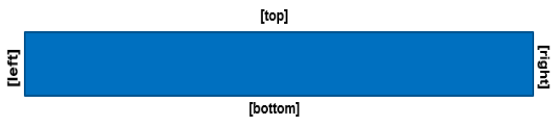


Figure 7. Boundary faces defined in Basilisk

Basilisk segregates the faces of the domain, which allows for the specification of boundary conditions. For example, the faces for a two-dimensional channel are shown in figure 7. The inlet is at the left boundary, and the outlet is at the right boundary. Boundary conditions were specified for velocity, pressure and tracer concentration. The full set of boundary conditions can be found in the Supplementary Information section 6. , alongside the Basilisk code.

A normalised unit number (dimensionless inlet velocity) can be set for constant velocity, or a function can be defined for more complex flows.

2.5 Boundary conditions: Oscillating flow

The boundary conditions can be set for each face of the large domain. Dirichlet or Neumann boundary conditions are set for each variable at the inlet (left face) and the outlet (right face); for example, on

velocity, pressure and tracer conditions. There are two types of flow simulated: constant velocity Poiseuille and oscillatory flow.

Oscillatory flow was achieved by imposing a time dependent Dirichlet boundary on the inlet face as such:

$$\texttt{u.n[left] = Dirichlet}\left[\frac{H}{2} + \frac{2H}{\pi}\sum_{1,\,n\,odd}^{N}\left(\frac{1}{n}\right)\sin\left(\frac{n\pi t}{L}\right)\right]$$

$$\frac{H}{2} + \frac{2H}{\pi}\sum_{1,\,n\,odd}^{N}\left(\frac{1}{n}\right)\sin\left(\frac{n\pi t}{L}\right) \tag{5}$$

Equation (5) denotes the Fourier series approximation of a periodic Heaviside function of amplitude H and frequency f. Adding more terms to the Fourier series will give more accurate Heaviside functions, however it was found that 20 terms gave computationally robust outputs with oscillations being negligible. Figure 8 shows the output for one such Heaviside function. Positive amplitude corresponds to 'forward' motion, whereas negative is 'reverse' motion. A net positive flow is achieved by making the upper bound higher than the lower bound.

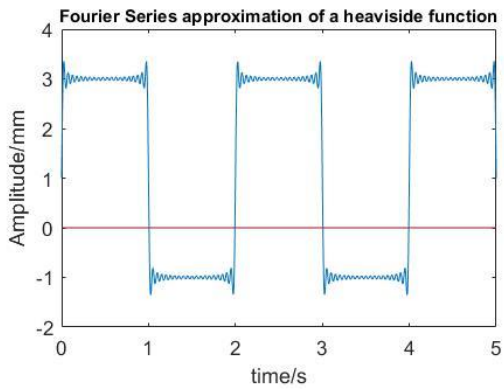


Figure 8. Modelling oscillatory flow using a Fourier series approximation of a Heaviside function

2.6 Output variables

The output variables are used to develop simulation videos, snapshots or output data for further analysis. At each timestep, the vorticity and the tracer concentration are gathered at every point in the domain. Furthermore, at the outlet face, the tracer concentration in every cell is normalised by the cell length. A summation of this gives the average tracer concentration, `f_mean`. This process is undertaken at every timestep.

2.7 Residence time distribution (RTD)

The mathematical treatment applied to the output variables follows that of McDonough (2019), though it is for two-dimensional simulations instead of three-dimensional experimental data. The exit age distribution, $E(t)$, is defined as the time spent by different elements in the reactor (Levenspiel, 1999). Given by equation (6), this is based on tracer concentration at the outlet. $C(t)$ is the average tracer concentration normalised by cell length. This is the variable `f_mean`.

$$E(t) = \frac{C(t)}{\int_0^\infty C(t)dt} \tag{6}$$

To facilitate comparison between different parameters and geometries, the exit age distribution is scaled using equations (7) and (8).

$$E(\theta) = \tau E(t)\,; \quad \theta = \frac{t}{\tau} \tag{7}$$

$$\tau = \frac{\sum tC(t)\Delta t}{\sum C(t)\Delta t} \tag{8}$$

τ signifies the residence time (the average time spent by an element of fluid), whereas θ denotes the dimensionless time. The corresponding $E(\theta)$ will be comparable across all reactor geometries and conditions. Strong plug-flow conditions have a 'narrow' distribution centred around $\theta = 1$, implying that most of the fluid exits the reactor at the given residence time. Poor plug-flow conditions, i.e. significant axial dispersion, are characterised by 'wider' distributions of $E(\theta)$ and the exit time is not necessarily centred around $\theta = 1$ (Li et al., 2016).

The tanks-in-series model quantifies the plug-flow performance, and stronger indicator than qualitatively considering the spread of exit-age distribution. The parameter N in the model signifies the equivalent number of tanks in series. Ideal plug flow is achieved when N tends towards infinity. Though $N \geq 10$ is generally accepted as the minimum for practical plug flow. N was determined by using equation (9) through mathematical solvers.

$$E(\theta) = \frac{N(N\theta)^{N-1}}{(N-1)!}e^{-N\theta} \tag{9}$$

Statistical moments, such as variance or skewness, were not considered as the tanks-in-series model gave a good quantifiable factor for performance. This is in general agreement with the analysis in McDonough (2019).

2.8 Diffusion

It should be noted that diffusion effects were not considered in the model. McDonough, (2019) suggested that the tanks-in-series parameter, N, can

be related to the Péclet number. The Péclet number is ratio of the advection rate to the diffusion rate. Larger Péclet numbers indicate better plug flow characteristics. This suggests, to optimise the system for PFR performance, diffusion transport has to be minimised in the design. Thus, not considering diffusion is a big limitation of the model. Implementing diffusion would require adding a second order derivative term to the Navier-Stokes equation. This was outside the scope of this project due to the computational and time constraints. Furthermore, no examples of diffusion flow problem have been previously implemented in Basilisk.

3. Results and discussion

3.1 The effect of oscillatory flow

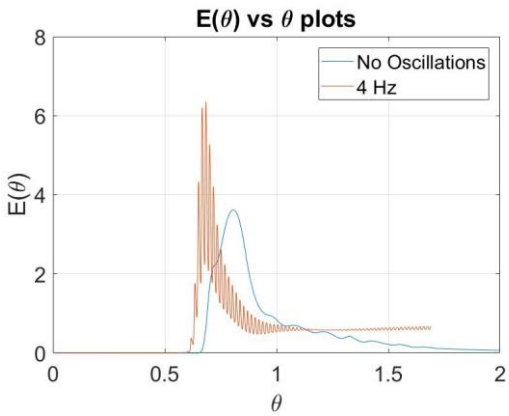


Figure 9. Exit distribution plot against θ, for constant-velocity flow and a 4 Hz oscillatory flow in a two-dimensional coil

Before the effects of oscillation frequency and amplitude could be studied, it was important to verify if the oscillatory flow really did improve plug flow performance compared to constant-velocity flow in a two-dimensional simulation. As per the experimental case of McDonough (2019), two main qualities were observed. First, as shown in figure 9 the oscillatory flow at 4 Hz has a much higher exit distribution, $E(\theta)$, peak in comparison to constant-velocity flow. Secondly, the shape of the 4 Hz curve is visibly narrower, exhibiting a more plug-flow-like behaviour. This effect is quantified below using a tanks-in-series model in figure 13. Both results are similar to the findings from the McDonough (2019) experiments. An interesting observation is that the exit distribution for the 4 Hz plot does not decrease to zero towards the end of the simulation. One reason for this could be the fact that tracer particles are being deposited at the walls of the tube due to the no-slip boundary condition and the absence of diffusion in the advection-only tracer-particle model.

The 4 Hz simulation exhibits a slightly better plug flow performance as seen from the residence time distribution (RTD) plot in figure 9. This is mainly due the interaction between the oscillatory fluid motion and the curvature of the reactor. The better plug flow behaviour due to oscillatory flow can be understood by analysing the flow in an OBR (Phan et al., 2010). The curvature of each coil in the tube acts as a baffle in an OBR, disrupting the boundary layer at the tube wall. The oscillation results in improved mixing through the formation of vortices. It is important to note that these vortices do not exhibit the same properties as Dean vortices, which can only be present in three dimensions.

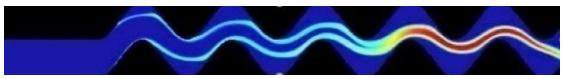


Figure 10. 2D simulation of a pulse tracer injection into a coiled tube with **no** oscillatory flow, showing a trail of tracer fluid lagging behind

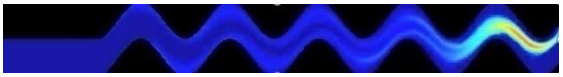


Figure 11. 2D simulation of a pulse tracer injection into a coiled tube with oscillatory flow of frequency 4 Hz, displaying a much better plug flow performance

This improved plug flow performance can also be seen in the simulations as shown by figure 10 and figure 11, both of which involve tracer sampling at the reactor outlet. The constant-velocity simulation produces a long and concentrated trail of the tracer compared to the simulation with 4 Hz oscillatory flow. However, in the oscillatory flow simulation the faint trail of tracer left behind is depositing to the walls of the tube, most likely due to the absence of tracer diffusivity.

3.2 The effect of varying frequencies

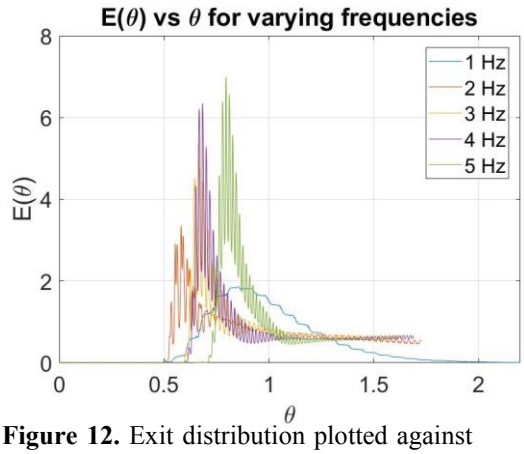


Figure 12. Exit distribution plotted against θ for various frequencies

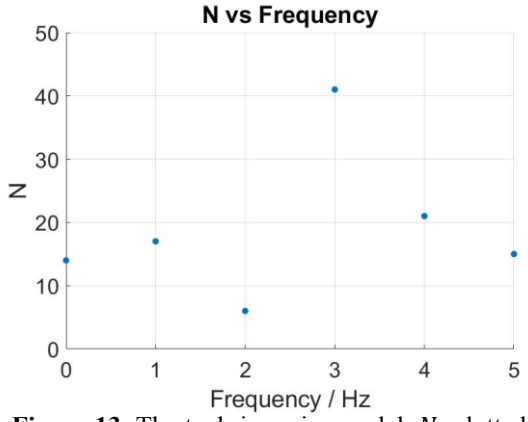


Figure 13. The tank-in-series model, N, plotted against frequency, showing a clear peak plug flow performance at 3 Hz

The effect of varying the oscillation frequency was investigated by holding the amplitude fixed at 2 mm. The resulting data was post-processed: $E(\theta)$ was plotted against θ as shown in figure 12. Furthermore, a graph of N against frequency was plotted as shown in figure 13 to better quantify the performance of plug flow in each simulation. As shown in figure 13, a frequency of 3 Hz generates the largest value of $N = 41$ and is therefore represents the best plug flow performance. This is similar to the results obtained by McDonough (2019) as shown in figure 2, where it was suggested that there is a 'switch-on' and 'switch-off' effect. This region corresponded to the $70 \leq \mathrm{De} \leq 200$, where secondary Dean instabilities are anticipated (McDonough et al., 2019). However, the results from this two-dimensional simulation cannot be directly compared to McDonough (2019) as Dean vortices are three-dimensional in nature. Figure 13 also displays a similar 'switch-on' and 'switch-off' effect, which suggests there is a two-dimensional analogy to secondary Dean instabilities. This is further backed by the fact that OBRs use baffles to generate vortices that promote radial mixing as mentioned in Section 1.1. This could be further investigated by simulating two-dimensional OBRs in Basilisk. It is also worth noting that a frequency of 2 Hz produced a worse N value when compared to constant-velocity flow.

3.3 The effect of varying amplitudes

Results from the McDonough (2019) studies highlighted how different oscillation amplitudes can significantly impact the plug flow performance. By fixing the frequency at 3 Hz and running two-dimensional simulations at various amplitudes, it can be seen from figure 14 that an amplitude of 2 mm generates the largest N value. This was in line with the results from McDonough (2019), also suggesting a presence of a two-dimensional equivalent to secondary Dean instabilities. However, it is not possible to directly compare results from a two-dimensional simulation to the three-dimensional experiment.

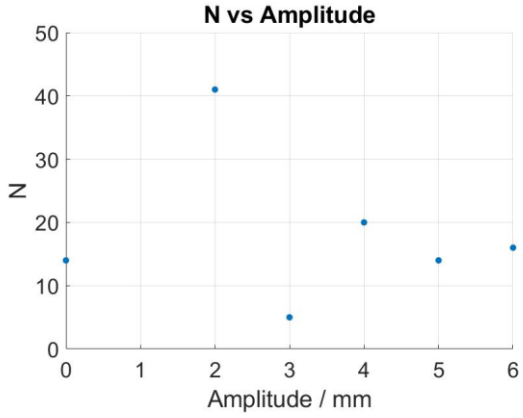


Figure 14. The tank-in-series model, N, plotted against amplitude, showing a clear peak plug flow performance at 2 mm

3.4 The effect of varying Reynolds Number

Setting the frequency of oscillations as 3 Hz and the amplitude as 2 mm, various different values of Reynolds number were tested. For $\mathrm{Re} \leq 100$, as Re increases, the plug flow performance decreases exponentially as shown in figure 15. It is important to note that reducing Re below 10 significantly increases the plug flow performance of the coil in the two-dimensional simulation. The N value increases from 41 at $\mathrm{Re} = 10$ to 59 when $\mathrm{Re} = 5$, representing a 44% increase. One of the main conclusions from McDonough (2019) stated that for $\mathrm{Re} < 10$, plug flow performance in coiled tubes can be significantly enhanced by oscillating the flow.

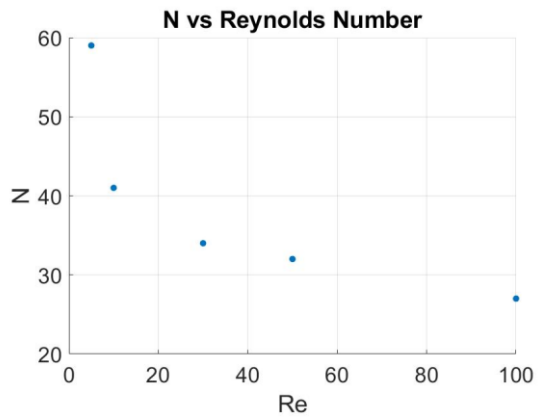


Figure 15. The tank-in-series model, N, plotted against Reynolds number, showing that plug flow performance is significantly enhanced for $\mathrm{Re} = 5$, in a coiled tube with oscillatory flow

3.5 Three-dimensional simulations

Due to time constraints, modelling of oscillatory flow was not achieved in three dimensions. However, as a proof of concept, the geometry was successfully implemented in the Basilisk framework and different sizes of coils were developed. Figure 16 displays the results of a simulation undertaken with fixed geometric parameters. Constant-velocity flow was simulated, and snapshots of profiles were extracted. Figure 17 and 18 shows the profiles of velocity magnitude and tracer concentration fields, at a defined plane. Constant-velocity flow took approximately one hour to execute on sixty computing threads and was found to be infeasible to run on a single computing thread. Therefore, it is reasonable to expect a timeframe of multiple hours for simple oscillating flow simulations.

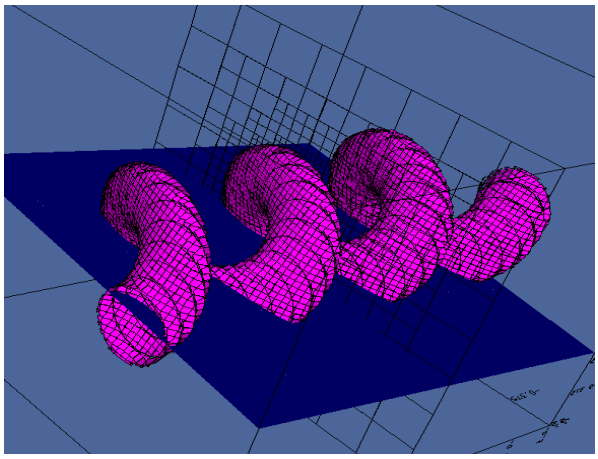
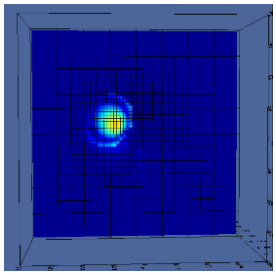


Figure 16. Three-dimensional constant velocity flow for parameter



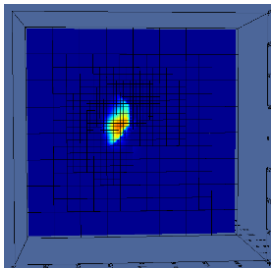


Figure 17 (left). Snapshot of velocity magnitude field at a defined two-dimensional projection

Figure 18 (right). Snapshot of tracer concentration field at a defined two-dimensional projection

3.6 Sensitivity analysis

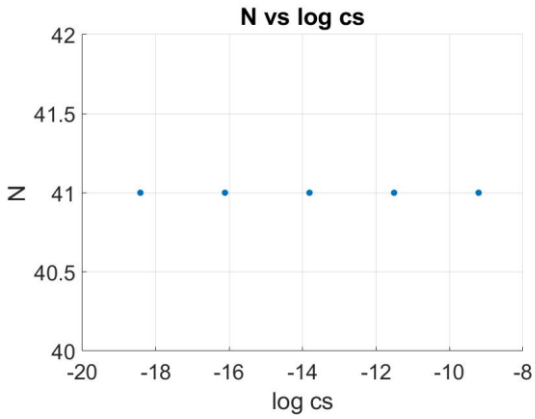


Figure 19. The tank-in-series model, N, plotted against `cs` tolerance, showing that the model is not sensitive to discretisation error

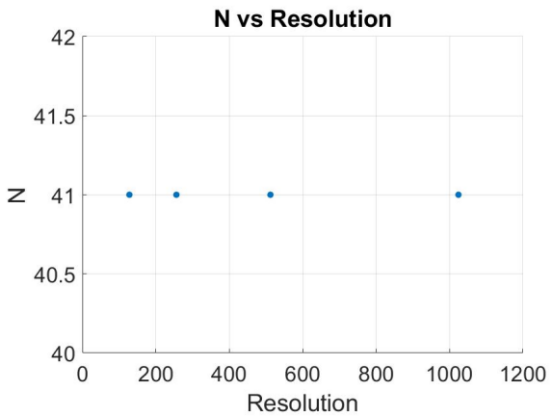


Figure 20. The tank-in-series model, N, plotted against `maxlevel`, showing that the model is not sensitive to `maxlevel` values

The discretisation error tolerance on the fluid volume fraction, `cs`, and simulation resolution, `maxlevel`, were both varied to check the sensitivity of the model. As shown in figure 19 and 20, neither variable affects the N value. This suggests that the Basilisk model configuration is computationally robust.

4. Conclusion

This project aimed to validate the experimental results from McDonough (2019). The objective was to model oscillatory flow in a helical reactor and to measure the improvement in plug flow performance compared to a constant-velocity flow. A general agreement was observed when comparing the N values at various frequencies and amplitudes to McDonough (2019). For example, the two-dimensional simulations from this project concluded 3 Hz and 2 mm both produced the best plug flow

behaviour in their respective tests which is in line with McDonough (2019). However, these results cannot be directly compared as the McDonough paper uses the idea of secondary Dean instabilities to explain their findings, whereas these instabilities are not present in two dimensions. Instead, it was hypothesised that the significant improvement in plug flow when comparing oscillatory and constant-velocity flow was based on the analogy of OBRs, where the curvature of each coil acts as a baffle to create vortices that promote radial mixing, hence improving the plug flow performance. It was also shown that the plug flow performance of a coiled tube is significantly enhanced at $\mathrm{Re} \leq 10$ when using an oscillatory velocity profile. Both the discretisation error, `cs`, and the resolution, `maxlevels`, were found to have no effect on the N value, and therefore the Basilisk model was considered computationally robust. Three-dimensional simulations were trialled, but additional research needs to be undertaken in order to model the oscillatory flow through a helix and to extract valuable data that can be used for validation against the experimental findings of McDonough (2019).

A key insight is that Basilisk has proven to be an excellent CFD program to model complex geometries and flows using implicit shape equations. The Basilisk adaptive mesh refinement tool was used for complex geometries that required zooming in on the walls to gain a better understanding of the flow. Most importantly, the prospect of easily integrating python scripts and machine learning algorithms with Basilisk code opens up a plethora of optimisation opportunities for the next phase of this project.

5. Outlook

The outlook for this project comprises of three key steps:

5.1 Confirming the fundamental validity of Basilisk results

Before proceeding onto three-dimensional simulations of a helical reactor, it would be worthwhile to investigate the validity of the fundamental Basilisk model. This would involve comparing a three-dimensional simulation of simpler geometries in Basilisk with real experimental data. It is recommended that this validation be done for a cylinder: a three-dimensional cylinder is very feasible to model with the current Basilisk code and useful RTD data can be easily extracted. This geometry is also easily reproducible in a lab, where residence time experiments can be performed by injecting a tracer fluid in a cylinder and measuring the exit distribution of the tracer. Other methods, to validate the Basilisk model include the experimental data (Hart et al., 2016) or using Laminar Flow Equation for analytical comparison (Fogler, 2016). By comparing the data from either of these resources with the Basilisk results, the underlying assumptions in the Basilisk code can be validated. Furthermore, to increase realism, diffusion would be implemented in Basilisk in the near future.

5.2 Validating McDonough (2019) results using 3D simulations and data

Due to the complexity of this project and the given time constraints, working three-dimensional simulations of oscillatory flow in a helix could not be obtained. One of the main conclusions of the McDonough 2019 report was that they required further CFD modelling in order to validate and to test the scalability of their findings. Therefore, the next phase of this project would involve working on the three-dimensional simulations and developing the oscillatory flow to work in a helix. One way to improve the present oscillatory flow model would be to replace the Fourier series approximation of the Heaviside function with modular arithmetic or some form of a periodic loop. This would eliminate the Gibbs phenomenon which is shown in figure 8. Furthermore, based on the current three-dimensional simulation, there are some bugs involved with the adaptive refinement near the walls of the helix. This leads to flow leaking outside the helix towards the end of the simulation. It is recommended that the adaptive refinement settings are tweaked in order to overcome this problem.

There are several ways to further develop this project. The next step would be to extract data for various operating conditions and reactor geometries and plot the relevant RTD curves. These results can then directly be compared to McDonough lab experiments and the performance of the Basilisk model can be evaluated. A key objective would be to visualise Dean vortices by looking at a cross section of a three-dimensional helix simulation. By running the simulation at a wide range of Dean numbers, both primary and secondary Dean vortices could be visualised.

5.3 Advanced multivariable optimisation using Machine Learning (ML)

As discussed earlier, the time constraints of the project only permitted single variable optimisation. The long-term outlook would involve implementing an advanced multi-variable optimisation algorithm

that considers various parameters such as oscillation frequency, amplitude, Reynolds number and various helical geometry parameters. Basilisk offers relatively easy integration with python and machine learning algorithms and this was one of the main reasons it was chosen over other commercially available CFD solvers. These optimisation techniques would then be applied to predict the optimum geometry for the helical reactor. This opens the scope for reactor design using CFD and machine learning instead of using the conventional design equations. It also allows the study of complex geometries such as a helix within a helix

Once these objectives have been achieved, the results will be passed back to the research team at Newcastle University who can print a three-dimensional prototype and test it experimentally. This opens up the possibility of rapid prototyping and testing of complex pilot scale reactors by significantly cutting down on fabrication time as well as the costs involved with reactor design.

References

Fogler, H. (2016). *Elements of chemical reaction engineering*. 5th ed. Upper Saddle River, N.J.: Prentice Hall PTR, pp.767 – 870.

Fuster D., Bague A., Boeck T., Le Moyne L., Leboissetier A., Popinet S., Ray P., Scardovelli S., and Zaleski S. (2009). Simulation of primary atomization with an octree adaptive mesh refinement and VOF method. *International Journal of Multiphase Flows*, 35(6):550–565, 2009.

Hart, J., Guymer, I., Sonnenwald, F. And Stovin, V. (2016). Residence Time Distributions for Turbulent, Critical, and Laminar Pipe Flow. *Journal of Hydraulic Engineering*, 142(9), p.04016024.

Harvey, A. And Phan, A. (2019). *Obrs*. [online] Pig.ncl.ac.uk. Available at: http://pig.ncl.ac.uk/new_page_2.htm [Accessed 07 Dec. 2019].

Law R., Ahmed S.M.R, Tang N., Phan A.N., Harvey A.P., Development of a more robust correlation for predicting heat transfer performance in oscillatory baffled reactors, *Chem. Eng. Process. Process Intensif.* 125 (2018) 133–138.

Levenspiel, O. (1999) Chemical Reaction Engineering. 3rd Edition. John Wiley & Sons, New York, 54

Li, S., Li L. And Xin, F. (2016). *Reaction Engineering*. 1st ed. Butterworth-Heinemann, pp.213-263.

Ligrani P.M. (1994) A Study of Dean Vortex Development and Structure in a Curved Rectangular Channel with Aspect Ratio of 40 at Dean Numbers up to 430 US Army Research Laboratory (ARL-CR-144) and Lewis Research Centre 4607

Masngut N., Harvey A.P., Ikwebe J., (2010) Potential uses of oscillatory based reactors for biofuel production, Biofuels 1 (4) 605–619.

Mcdonough, J., Murta, S., Law, R. And Harvey, A. (2019). Oscillatory fluid motion unlocks plug flow operation in helical tube reactors at lower Reynolds numbers (Re ≤ 10). *Chemical Engineering Journal*, 358, pp.643-657.

Nivedita, N., Ligrani, P. And Papautsky, I. (2017). Dean Flow Dynamics in Low-Aspect Ratio Spiral Microchannels. *Scientific Reports*, 7(1).

Parida D., Serra C.A., Garg D.K., Hoarau Y., Bally F., Muller R., Bouquey M., (2014) Coil flow inversion as a route to control polymerization in microreactors, Macromoleculaes 47 (10) 3282–3287

Phan, A. And Harvey, A. (2010). Development and evaluation of novel designs of continuous mesoscale oscillatory baffled reactors. *Chemical Engineering Journal*, 159(1-3), pp.212-219.

Popinet S. (2016) . Basilisk: simple abstractions for octree-adaptive scheme. In *SIAM conference on Parallel Processing for Scientific Computing, April 12-15 2016, Paris*, 2016.

Popinet, S. (2019). *Basilisk – Basilisk*. [online] Basilisk.fr. Available at: http://basilisk.fr/ [Accessed 1 Nov. 2019]

Reay, D., Ramshaw, C. And Harvey, A. (2013). *Process intensification*. 2nd ed. Butterworth-Heinemann, pp.121-204.

Sharma A.K., Agarwal H., Pathak M., Nigam K.D.P, Rathore A.S., (2016) Continuous refolding of a biotech therapeutic in a novel Coiled Flow Inverter Reactor, Chem. Eng. Sci. 140153–160

Thoraval M.-J., Takehara K., Etoh T. G., Popinet S., Ray P., Josserand C., Zaleski S. And Thoroddsen S. T. (2012) von karman vortex street within an impacting drop. *Physical Review Letters*, 108(26):264506.

Van Hooft, J.A., Popinet, S., van Heerwaarden, C.C. (2018). Towards Adaptive Grids for Atmospheric Boundary-Layer Simulations. *Boundary-Layer Meteorol.* 167, 421–443 (2018) doi:10.1007/s10546-018-0335-9

Hydrogen Infrastructure Spatio-Temporal Optimisation for the UK's Heat Network

Hassan Sherbaz, Klement Tam

Abstract

Heating for residential and industrial buildings releases 110Mt of CO_2 annually. 85% of heating comes from natural gas, which is not compatible with the Committee on Climate Change's proposed target of net zero greenhouse gas emissions by 2050. Hydrogen can be a low-carbon replacement to natural gas due to it's storage flexibility and minimal transmission losses. To identify the key opportunities and challenges in a hydrogen-for-heat network, the Hydrogen Infrastructure Spatio-Temporal Optimisation (HISTO) model was developed. This illustrates a gradual transition from natural gas to hydrogen from 2020 to 2050. HISTO also studies the intraday and interseasonal demand fluctuations. A Pareto frontier was developed between the Total Discounted Cost against an Emission Budget Target by 2050. Each data point provides a spatio-temporal optimisation of the UK grid. Different scenarios were studied. This highlighted that using a 20mol% blend of hydrogen with methane would provide a minimum of 16% emission cut, without the need for appliance or transmission overhaul, providing a "low-regret" solution for 2030. Underground storage was also imperative as they provided a guarantee of supply even if there was a large reliance on intermittent sources of production i.e. wind and electrolysers. Finally, allowing the wind generators to sell electricity to the grid and the electrolysers enhanced the uptake of electrolysers, therefore some integration of the heat and electricity grid can be economical. Such results can be used to better inform policy makers, local authorities, utility providers and investors.

Keywords - Spatio-Temporal, Hydrogen for heat, MILP, Supply chain optimisation

1. Introduction

In a typical year, the UK consumes 328TWh of heat energy with a peak demand of 159GW [1]. This results in 110Mt of CO_2 emissions, 30% of the total annual emissions. 85% of this heating is supplied from natural gas, which emits 0.184 kg CO_2/kWh [2]. Currently, 56% of the natural gas supply for electricity generation and heating is imported, most of which is via European pipelines from Russia. As North Sea gas fields start to deplete, dependence on imports for natural gas will increase, thus increasing political dependence to meet energy requirements. Despite this, household consumption of electricity and heat has fallen by 17% in the last decade due to enhancements in building efficiency and insulation which is significant considering the number of households has increased. The Committee on Climate Change (CCC) published a report which replaces the 80% Greenhouse Gas (GHGs) reduction target by 2050, where they now propose a 100% reduction [3]. Consequently, radical action must be taken to tackle all emissions sources within the UK. This means that alternatives to natural gas for heating need to be implemented on a national scale within the next 30 years. The CCC has outlined several "low-regret" approaches which can be a feature in a number of different decarbonisation targets [4]. These include:

1) Continuing to improve energy efficiency in existing and new buildings
2) Low carbon heat networks in heat dense areas. This can be through heat recovery from Large Biomass, Municipal Waste, and other Industrial activities.
3) Heat pumps in the half of the 4 million homes not connected to the mains gas network

This, however, does not tackle the 20 million gas-heated homes in areas where heat networks are not cost effective. The decarbonisation strategies for such buildings include:

1) Electrification using highly efficient heat pumps and storage heaters for smaller properties.
2) Shifting to a green gas supply such as low-carbon hydrogen
3) Hybrid heat pumps where heat pumps meet most of the demand and hydrogen or a blend can meet the peaks.

Imperial College London was commissioned to provide a study on this which showed all three pathways had annual costs within 10% of each other [3]. This was in line with estimates by Element Energy Ltd . This suggests that for a complete decarbonisation, a combination of all 3 pathways could be deployed. Such costs also meet budgets of less than 1% GDP as proposed by CCC [4]. These findings show that to cut emissions, a low-carbon gas is necessary to deploy serious decarbonisation pathways. The main reasons for why hydrogen is being considered as a potential low-carbon gas are as follows:

1) Transmission losses are minimal relative to other technologies. For electricity, 9% is lost from production. District heat and hot water networks are even more inefficient and costly to transport over long distances. Consequently, hydrogen has the potential for centralised production, transportation, storage and local conversion to heat.
2) Hydrogen, like natural gas is easily stored. On the other hand, electrification of heat would not provide the same flexibility due to the limitations of electric batteries. Currently, batteries would incur large losses if they were used to meet interseasonal demand and are also too small in scale. Direct thermal storage for long periods is also difficult due the required insulation of the storage device.

This is too expensive to consider as a national-scale solution.

Because of this, hydrogen is being considered seriously as a low-carbon gas. Detailed planning is required to inform investors and policymakers as well as to set up pilot projects. This is because several challenges currently exist for hydrogen-for-heating. The main production methods for hydrogen are electrolysis, methane reforming and gasification (Coal or Biomass). For electrolysis, the electricity supplied must come from renewable sources such as wind to ensure that zero-carbon hydrogen is produced. Wind is an intermittent source of electricity, which means the production is not guaranteed. Methane reforming produces CO_2 and hydrogen, which therefore requires Carbon Capture and Storage (CCS). Similarly, gasification produces CO_2 and requires CCS. Ultimately, this raises the problem of which production facilities to use, and in which locations considering varying demand, underground formation for storage, and area availability.

2. Literature Review

Commissioned industry reports provide a good basis to define the problem. Several models have been built to analyse the feasibility of adopting hydrogen as a fuel. Equilibrium energy models such as MARKAL/TIMES match supply and demand over a given planning period [5]. However, this does not include any spatial resolution, and interseasonal and intraday supply and demand. Supply chain models for hydrogen have been developed in the context of the mobility sector and can provide more in-depth insights. Such models can still be adapted for a hydrogen for heating supply chain model.

2.1. Commissioned Reports

One of the key strategies in the *Decarbonisation of Heat* report [6] is providing a low-carbon solution for buildings that are on the gas grid and not on heat networks. This constitutes most of the current heat demand. The *Next Steps on UK Heat policy* report [3], recommended that the 20 million homes currently connected to the central natural gas grid could be decarbonised by using hydrogen.

Consequently, there has been a greater initiative to test the feasibility of using hydrogen since 2016. The BEIS hydrogen supply chain evidence base [7] provides a basis for the hydrogen production technologies with their associated costs and learning rates up to 2050.

The H21 report [8] provided a localised study of Leeds' heating supply and found that by leveraging the existing network, a natural gas to hydrogen conversion is feasible. The end result showed that 4 steam methane reformers were constructed with a total capacity of 1025 MW_{HHV}, with interseasonal and intraday demand balanced using salt caverns in Teesside.

In terms of live pilot projects, HyDeploy [9] started a live test in Keele University (Phase 1) where a 20mol% hydrogen blend with natural gas is being tested for heating. This has now been proven to be safe without the need to replace pipelines or boilers. Phase 2 and 3 will see a wider gas blend roll out to different residential properties in the North East. The hydrogen is being produced using electrolysis. Meanwhile in the North West, HyNet is looking to produce hydrogen through Auto-Thermal Reformers (ATRs) whilst capturing the CO_2 and storing it in Liverpool Bay's oil and gas fields. The aim is to build 890 MW of hydrogen output capacity from ATRs with a 93% CO_2 capture rate, where the hydrogen can be injected to the Local Transmission Network (LTS). However, the *Element Energy Ltd* report [7] states that current ATR designs are more suited to Syngas production and must be re-optimised for pure hydrogen production. For that reason, previous modelling work has focused on using SMRs as it is a more established design.

2.2. Mobility Sector Spatial Optimisation

Much of the early research in hydrogen supply chain optimisation was focused on the mobility sector. One of the first models which applied spatial optimisation was the Almansoori model [10]. It developed a future hydrogen supply chain for fuel cells in the mobility sector within the UK. The model divided the UK into 34 spatial zones where hydrogen can be transferred via road tankers to fuelling stations.

Several modifications have built on top of this spatial formulation. Moreno added constraints that consider carbon (for CCS) and hydrogen pipelines for distribution across the UK [11]. Dayhim applied a demand forecast on New Jersey by analysing how uncertainty in socio-economic factors affected this network [12]. Kim and Moon considered safety as well as profit, and allowed for hydrogen production through wind turbines with electrolysers [13]. Sabio's model took a holistic view of minimising emissions by carrying out a full Life Cycle Assessment [14].

The main drawback of these models is that they are "snapshot" models, i.e. only the spatial optimisation of the network are considered at a single point in time. The interseasonal and intraday demand shifts were not accounted for, prohibiting a comprehensive modelling of operating an energy network. Storage is often considered as a static parameter expressed as a certain number of days of demand stored, e.g. 10 or 20 days. This level of analysis will not exemplify the benefits of storing and transmitting hydrogen at different times of the year and day, especially when it is produced through variable renewable sources. Except Kim and Moon, all the models discussed do not consider wind generation to produce hydrogen. The Kim and Moon paper takes a static average production rate which underestimates the intermittency challenges of relying on renewable resources.

2.3. Spatio-Temporal Optimisation

Newer models achieved more comprehensive studies of potential hydrogen supply chains by adding a time-dependent element.

Lahnaoiu developed a model optimisation on North Rhine-Westphalia, Germany, where hydrogen is produced from excess wind electricity [15]. This was applied to the mobility sector, where optimal truck paths to the fuelling stations were identified.

Reus's model highlighted the challenges of matching intermittent supply of hydrogen through wind and the varying demand of in Germany's mobility sector [16]. It provided a detailed study of seasonal storage and performed wind turbine site suitability modelling using Geographic Information System (GIS) software.

Samsatli's Value Web Model (VWM) took a focused view on hydrogen-for-heating in the UK [17]. Similar to Reus, it utilised GIS modelling to assist in the planning of building new wind turbines. The model studied a combined network situation where heat and electricity grids are integrated which will provide added synergies and a smarter network. However, the end recommendation of using 80% electric heating and 20% hydrogen for heating might be more difficult in practice. The heating demand peaks are 5 times greater than the electricity peaks, adding 300GW [6]. This will require significant reinforcement to the electricity grid and such associated costs are not considered in the VWM. Consequently, an 80% direct electric heating future contradicts the commissioned reports discussed above.

2.4. Contributions Of Work

The discrepancy between the commissioned reports analysing UK's heating demands and current spatio-temporal models remains significant. This work aims to provide insights which can bridge this gap further and better inform policy makers and investors. In particular, none of the previous models considered a gradual transition over time from the current predominantly-used natural gas to a future hydrogen infrastructure network, specifically for the heating industry. By building a novel spatio-temporal optimisation model for a hydrogen supply chain without reliance of an extensive integration with the electricity grid, it will provide a standalone study for how this transition can occur. This model can then be further modified accordingly to test potential real cases of hydrogen deployment such as using blends, wind outages and minimal grid integration.

3. Methods

3.1. Spatio-Temporal Optimisation

In this study, UK is spatially distributed into 12 zones based on the regions which the Department for Business, Energy and Industrial Strategy (BEIS) collects its gas consumption statistics. Temporal representations are important to monitor network movements across the 30-year horizon from 2020 to 2050. 5-year planning period intervals are considered for investment decisions. Each planning period interval comprises the 4 seasons (Winter, Spring, Autumn, and Summer) and each season consist of 4 times of the day (Day, Evening Peak, Late Evening, Night). Altogether, these form 16 non-uniform time slices per 5-year planning period.

3.2. Infrastructure Network

Fig. 1 depicts the hydrogen infrastructure network diagram across the three components of the supply chain considered – production, storage and transmission. Hydrogen can be generated within the zone either by steam methane reforming with 90% carbon capture storage (SMRCCS), onshore wind turbines with proton exchange membrane electrolysers (ONWTE), or by offshore wind turbines with PEM electrolysers in the waters for coastal zones (OFFWTE). It is assumed that the wind turbine and electrolyser are built in the same zone. The produced hydrogen can then be stored in underground salt caverns (if present in the zone) or pressurised vessels which can be built in 3 different sizes. Transmission within neighbouring zones is viable through pipelines or trucks. Together with the natural gas network assumed to be already present in all zones, the hydrogen generated in a zone or received from other zones, can then be used to satisfy local heat demands.

The technologies chosen for this study are those that have been well supported by previous work on hydrogen supply chain and are financially practical. For example, PEM electrolysers were favoured since it has been scientifically tested and has a higher learning rate on capital cost than the more mature alkaline electrolysers [7]. Other technologies such as gasification techniques were not considered as the design is still in the early stages for pure hydrogen production and has excessive predicted costs.

3.3. Hydrogen Infrastructure Spatio-Temporal Optimisation Model (HISTO) Mathematical Formulations

The novel Hydrogen Infrastructure Spatio-Temporal Optimisation model (HISTO) is a mixed-integer linear program (MILP). Optimal solutions for the investment and operation of the hydrogen supply chain infrastructure (including transmission and managing of storage inventory) were found by implementing the proposed algorithm using the CPLEX solver accessed via the GAMS modelling tool. The nomenclature and complete set of formulations are found in the Supplementary Material.

3.3.1. Material balance: The flows across production, transmission and storage in each zone and time slice are represented by the material balance in Eq. (1), which is an inequality constraint that balances these flows. The demand for each zone, is obtained from current gas consumption statistics for 12 zones in the UK [18]. Allocated demand for residential and industrial properties are aggregated then split into 16 time slices. The total heat demand was converted by assuming a natural gas boiler efficiency of 80%. Hydrogen boiler efficiency was assumed at 90% [19]. A conservative assumption was made that the annual heating demand will remain the same as there could be a trade-off between improved building efficiency and more buildings constructed.

$$\sum_{z't}(QT_{zz'tyh} - QT_{zz'tyh}) + \sum_{s}(QSout_{zsyh} - QSin_{zsyh}) + \sum_{p} PRD_{zpyh} - Demand_{zh} = 0, \forall z, y, h \quad (1)$$

3.3.2. Production: The hydrogen production rate in each zone and time slice is bounded by the installed capacity of production technologies in the zone, which can increase across planning periods as seen in Eqs. (2) and (3). A production

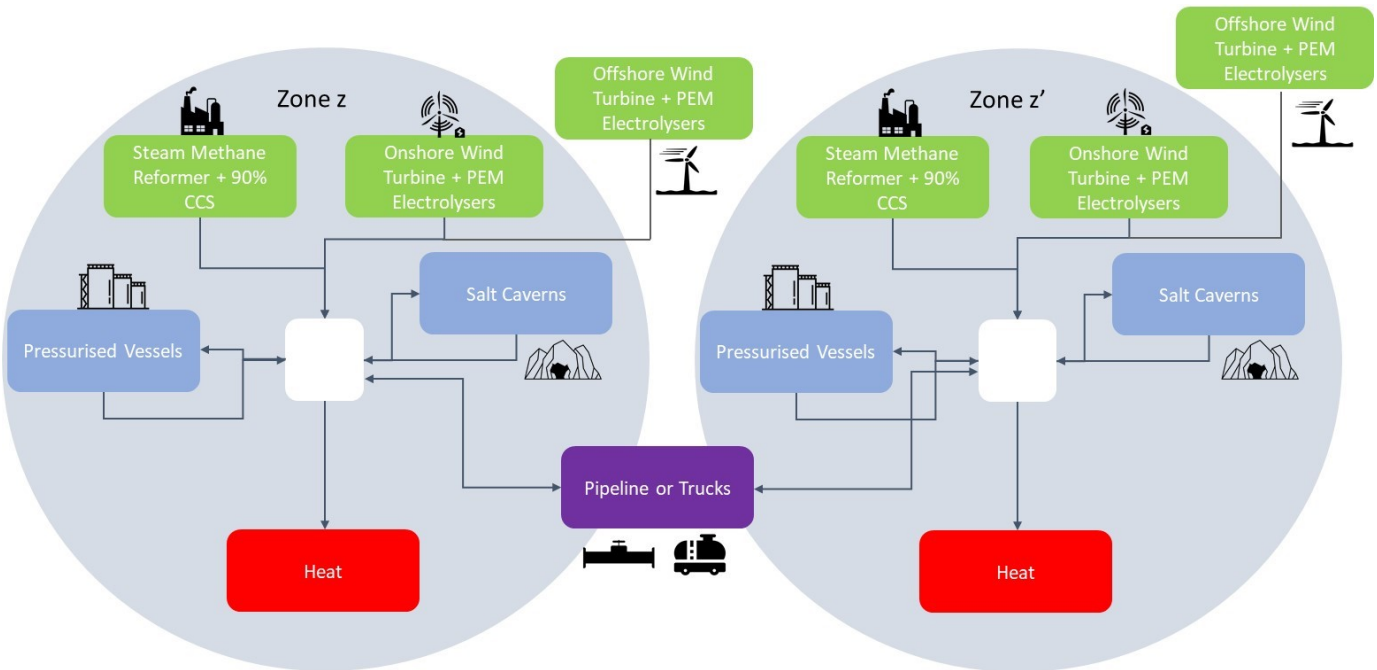


Fig. 1: Hydrogen infrastructure network diagram showing all potential supply chain pathways

factor is included in Eq. (2), which accounts for the varying wind speed profiles in different seasons and times of the day as obtained from Sinden (2007) [20]. Since installing onshore wind farms may spark concerns on land availability, the number of new onshore wind turbines installed in each zone are restricted using Eq. (5). The factor $2\sqrt{3}$ assumes that the turbines are arranged on a hexagonal grid. Data for the total available area for wind turbines in each zone are adapted from Samsatli et al. (2016) [21] which used GIS mapping.

$$Producemin_p \cdot PRDFactor_{ph} \cdot PRDCAP_{pzy} \leq PRD_{pzyh} \leq PRDFactor_{ph} \cdot PRDCAP_{pzy}, \forall z, p, y, h \quad (2)$$

$$PRDCAP_{pzy} = PRDCAP_{pzy-1} + PRDCAPI_{pzy}, \quad \forall z, p, y \geq 2 \quad (3)$$

$$PCAP_p^{min} \cdot NPI_{pzy} \leq PRDCAPI_{pzy} \leq PCAP_p^{max} \cdot NPI_{pzy}, \forall z, p, y \quad (4)$$

$$\sum_y 2\sqrt{3}(5R^w)^2 \cdot NPI_{pzy} \leq PCAP_p^{max} \cdot NPI_{pzy}, \quad \forall z, p = ONWTE \quad (5)$$

3.3.3. Storage: A material balance is used in Eq. (6) for the amount of hydrogen that is stored between timeslices. The stored amount is constrained by the storage capacity in Eq. (7) which similar to production, can increase every planning period for that of pressurised vessels. Binary variables in Eqs. (8) – (10) enforces that hydrogen is either added or removed from the storage facility in each timeslice and not both simultaneously. Flowrates into and out of storage facilities are also bounded by maximum deliverability and injection rates.

$$Store_{szyh} = Store_{szy,h-1} + (QSin_{szy,h-1} - QSout_{szy,h-1}) \cdot Duration_h, \forall z, s, y, h \geq 2 \quad (6)$$

$$Store_{szyh} \leq STRCAP_{s,z,y}, \forall z, s, y, h \quad (7)$$

$$SIN_{szyh} + SOUT_{szyh} \leq 1, \forall z, s, y, h \quad (8)$$

$$\frac{SIN_{szyh}}{M_a} \leq QSin_{szyh} \leq SIN_{szyh} \cdot MaxInj_{sz}, \forall z, s, y, h \quad (9)$$

$$\frac{SOUT_{szyh}}{M_a} \leq QSout_{szyh} \leq SOUT_{szyh} \cdot MaxDel_{sz}, \quad \forall z, s, y, h \quad (10)$$

3.3.4. Transmission: Hydrogen can be transmitted between zones via pipelines or trucks. Eq. (11) - (12) ensures unidirectional pipeline flow between 2 zones. The big-M method was used in Eq. (12) to determine whether there is hydrogen flow between 2 zones. In Eq. (13), pipeline flows are bounded by the total pipeline capacity, which is attained from assuming a constant pipeline diameter of 10 inches. Eq. (14) allows for additional pipelines to be built in each planning period.

$$X_{zz'yh} + X_{z'zyh} \leq 1, \forall z, z', y, h \quad (11)$$

$$\frac{X_{zz'yh}}{M_a} \leq QT_{zz'yh}^{pipe} \leq X_{zz'yh} \cdot M_a, \forall z, z', y, h \quad (12)$$

$$QT_{zz'yh}^{pipe} \leq NT_{zz'y}^{pipe} \cdot QTCAP_{zz'}^{pipe}, \forall z, z', y, h \quad (13)$$

$$NT_{zz'y}^{pipe} = NT_{zz',y-1}^{pipe} + NTI_{zz'y}^{pipe}, \forall z, z', y \geq 2 \quad (14)$$

Eq. (15) calculates the flowrate transmitted via trucks, taking into account factors such as loading and unloading time, average truck speed, road distances etc. Eq. (16) bounds

the amount of hydrogen transported while Eq. (17) allows additional truck units.

$$QT^{truck}_{zz'yh} = \frac{TruckOpTime}{\frac{2RoadDistance_{zz'}}{TruckAvgSpeed} + LUTime} \cdot \frac{VT^{Truck}_{zz'yh}}{Duration_h}, \quad \forall z, z', y, h \tag{15}$$

$$0 \leq \sum_{z'} VT^{Truck}_{zz'yh} \leq NT^{truck}_{zy} \cdot QTCAP^{truck}, \forall z, y, h \tag{16}$$

$$NT^{truck}_{zy} = NT^{truck}_{z,y-1} + NTI^{truck}_{zy}, \forall z, y \geq 2 \tag{17}$$

3.3.5. Emission budget: The final category of constraints includes the emission budget constraint. As seen from Eq. (18), an inequality constraint is placed on the amount of CO_2 that can be emitted from the production technologies across the network. The right side of the equation shows the parameter EBT, which is a set emission budget target at 2050, i.e. a proportion of current emissions from natural gas that is allowed. A linear decrease of the permissible emissions is assumed over the planning periods (2020-2050) to achieve this set carbon emission budget target.

$$\sum_{p,z,h} CO_2Emission_p \cdot PRD_{zpyh} \cdot Duration_h \leq CurrentEmission \cdot EBT \cdot \frac{y - 2015}{35}, \forall y \tag{18}$$

3.3.6. Objective function: The objective function minimizes the total discounted costs, which takes into account the investments and costs of operation of the entire hydrogen network in each planning period. The costs and learning rates for production, storage and transmission were obtained from the Element Energy Ltd, and the H21 reports [22], [8].

$$\min_{s.t.} TDC = \sum_{y} \frac{TC_y}{(1+r)^{y-1}}, \ where \tag{19}$$

$$TC_y = CAPEXP_y + OPEXP_y + CAPEXS_y + OPEXS_y + CAPEXT_y + OPEXT_y \tag{20}$$

4. RESULTS AND DISCUSSION

The model was implemented for a base scenario which examines a bi-criterion optimisation approach in the design and planning of the hydrogen network. In addition, the optimisation model was configured for 3 case studies which provides further depth into the proposal of using a hydrogen network for heating. The base case and case studies are as follows:

- Base case - examining how TDC varies as the EBT parameter is tightened, and probing into the optimised spatio-temporal solution across the supply chain for a specific EBT.
- Case 1 - investigating how a live pilot project, in particular HyDeploy, would work on a large scale across the entire UK.
- Case 2 - exploring the impacts of wind outages in deep decarbonisation scenarios, where there is a high adoption of wind turbines with electrolysers.
- Case 3 - analysing how selling part of the wind turbine installed capacity to the electricity grid can affect adoption of electrolysers.

All the MILP problems were implemented in GAMS and solved using the CPLEX solver on a High Performance Computer. For each EBT, the simulation consists of 149,666 variables (of which 33,972 are integer variables) and 178,606 constraints. The total CPU time taken varied across different scenarios, with the longest (EBT = 0.1) taking 21 minutes. An optimality tolerance of 0.1% was chosen (i.e. the best integer objective must be within 0.1% of the objective of a fully-relaxed problem) to ensure a sufficient level of accuracy.

4.1. Base Scenario

4.1.1. Pareto solution: By introducing a *LOOP* function, the MILP optimisation problem was solved multiple times, with each consecutive iteration having a different EBT and thus different allowable emissions in the same planning period. This generated a Pareto frontier as shown in Fig. (2), illustrating the trade-off between TDC and EBT. Fig. (3) shows the accumulated total heating energy supplied from 2020-2050 with different values of EBT.

The frontier can be segmented into 4 categories: A) A complete natural gas grid when there is no limit on carbon emissions; B) Linear uptake of SMRCCS due to its steady adoption against natural gas; C) Emergence and uptake of ONWTE as the EBT becomes too tight for SMRCCS to exist as the only hydrogen production technology; D) When ONWTE exists as the only hydrogen production technology for complete decarbonisation by 2050.

ONWTE only emerges when the EBT is under 0.2 due to its higher costs (driven by electrolysers) relative to SMRCCS, despite the former having a more favourable learning rate on capital costs. Consequently, a steep increase in TDC is incurred as more ONWTE is adopted, making deep depcarbonisation scenarios less financially feasible. The most expensive option of OFFWTE is not adopted throughout, which shows that there would be sufficient area for onshore wind turbines if a hydrogen network were to be implemented.

It is worth noting that if ATRs were considered, they would replace all SMRs as they allow for a higher capture rate (95%) at a similar costs. However, they were not included as functional designs optimal for hydrogen production are not as mature as SMRs [7].

4.1.2. Model validation: To achieve complete decarbonisation, i.e. EBT of 0, this incurs an additional £15.6bn relative to point A on Fig. (2). This amounts to a cost of avoided emissions of £56.5/tCO_2 saved and a levelised cost of heat (LCOH) of £58.1/MWh.

The Value Web Model predicts that for a 82.6Mt CO_2 annual saving, it would cost £8.54/tCO_2 with a LCOH of £38.28/MWh (current nominal heat price £33.3/MWh) [23]. The HISTO model requires a much more expensive cost of emission avoidance £39.5/tCO_2 and a LCOH £45.5/MWh . On the other hand, the H21 report states that the 4 SMRCCS units for Leeds would require £291.4/tCO_2 for emission avoidance [8]. As the scale of the HISTO model is much greater than the H21 study, a more connected network would provide cost savings not found in a localised study. However, the H21 reports provides greater granularity in the cost structure,

which makes it more conservative than the HISTO model. On the contrary, the VWM allows for 80% of heating to be electrified, saving costs on SMRs and PEM electrolysers whilst not considering the cost of electric grid reinforcement required. Consequently, the VWM could potentially be underestimating the feasibility of its optimal solution.

A technology abatement curve done by McKinsey includes several technologies which would cost below £50.5/tCO_2. Similar to the HISTO model, it calculates the abatement rates assuming aggressive uptake of a particular technology. Currently, complete decarbonisation in the HISTO model exceeds this by £6/ tCO2. Despite this, EBT = 0.1 has comparable abatement costs to other technologies such as Gas and Coal plant CCS retrofit [24].

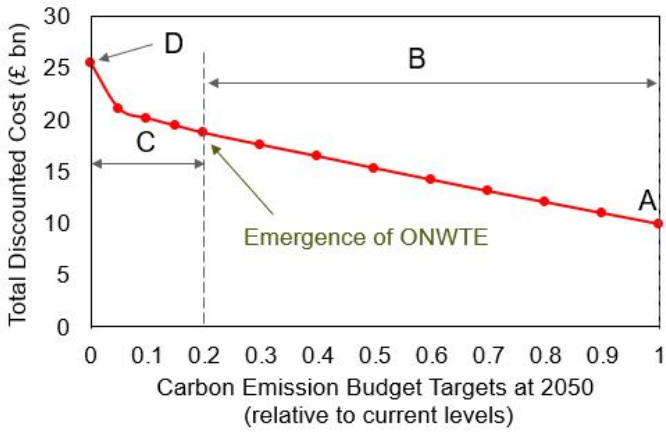


Fig. 2: Pareto solution of TDC against EBT

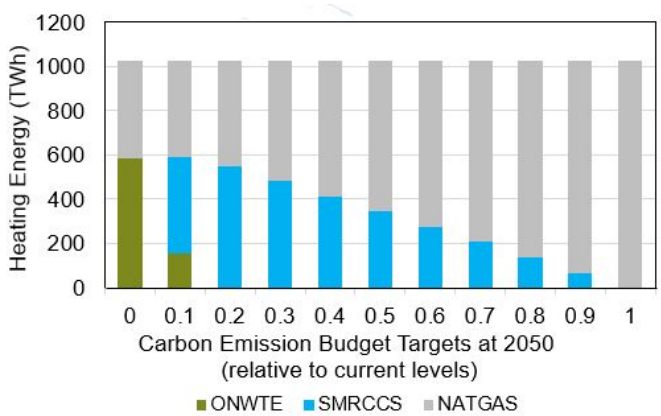


Fig. 3: Total heating energy supplied from 2020 to 2050

4.1.3. Hydrogen network transition: Fig. (4) dives into the hydrogen network transition over the time horizon when the EBT is set at 0.1. This has an LCOH of £48.6/MWh which is a 46% increase from the current nominal price. This would have to be subsidised by the government, but the end customers will also face a large increase. Only 3 planning periods are shown. The complete network over time can be found in the Supplementary Material. In the first planning period (not shown here), the model builds SMRCCS in all zones except for Z7 and Z8 (Inner and Outer London), which has been prohibited due to limitations of major industrial activities in the capital. As time progresses towards 2030, more SMRCCS units are installed. Zones with underground salt caverns (Z2, Z3, Z9) are key locations for hydrogen production units.

As seen from Fig. (4a), the pipeline network shows a disconnect between North and South England until 2035. As the transition from natural gas to hydrogen extends due to emission constraints, the salt caverns' capacity in Z9 is not large enough to provide the flexibility in surrounding regions (Z7, Z8, Z9 and Z10). Thus, the 2 pipelines connecting Z4 with Z6 and Z6 with Z8 introduced in Fig. (4b) are crucial to enable the Southern zones to have access to the much larger salt caverns in Z3 (Yorkshire). By 2035, the hydrogen pipeline network connects all the zones accordingly.

From 2040 onwards, the emission budget becomes too small for only SMRCCS to be built, explaining the large uptake of ONWTE. The final network in 2050 is shown in Fig. (4c), where 52% of heating comes from ONWTE distributed across UK depending on area and wind availability of each zone.

4.1.4. Interseasonal storage: No pressurised vessels are installed throughout the years, which demonstrates the high costs of the technology. However, underground salt caverns play a huge role in the hydrogen network. Fig. (5) displays the optimal storage inventory over a typical year. In all the 3 zones with salt caverns, hydrogen is accumulated in the summer to reach maximum capacity in autumn, before distributed to the surrounding zones during winter to fulfill the higher heating demands.

4.1.5. Intraday storage: Fig. (6) shows the total net flow into all storage units at different times of the day during the 4 seasons. During spring, there is a varied flow of hydrogen at different times of the day. A positive net flow into storage units is seen throughout the summer time slices, explaining the steep accumulation of hydrogen inventory in Fig. (5). Storage inventory movements are relatively neutral during autumn. Key periods during winter include the day (07:00h-17:00h) and evening peaks (17:00h-20:00h), where the high demand requires a large outflow of hydrogen from the storage units.

4.2. Case 1 - HyDeploy

HyDeploy is a pilot scale project testing the feasibility and safety of using a 20mol% hydrogen blend with natural gas. Phase 1 became live in Autumn 2019 where the blend was used to heat Keele University. It is expected that Phase 2 and 3 will lead to supplying heat for residential homes in the North East area. The significance of this blend is that it does not require appliance or pipeline overhaul. The existing infrastructure can be utilised. Consequently, the base model was augmented to analyse the emission cut, TDC and the optimal network which can be achieved if such a blend was used on a national scale. One of the safety issues associated with using a blend is that hydrogen's flammability limits are wide (between 4 vol% to 75 vol%). This is why the HyDeploy project is testing a wide variety of end user scenarios in Phase 2 and 3.

Fig. (7) illustrates that a 16% emission cut can be achieved with a 20mol% hydrogen blend. The optimised solution will

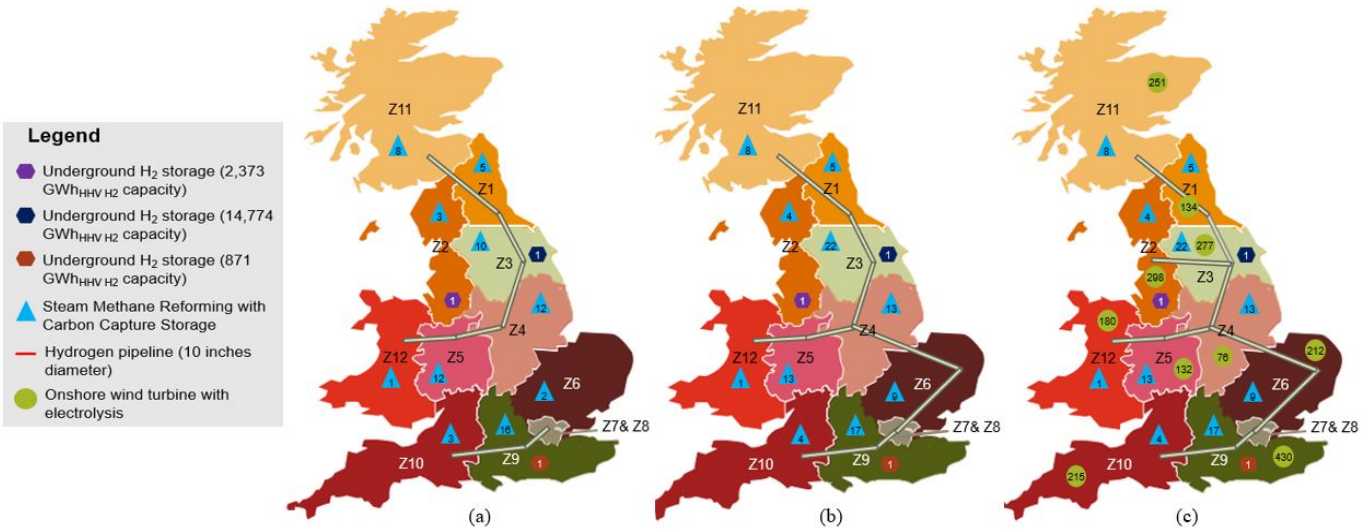


Fig. 4: Optimal hydrogen infrastructure network in the base case for 0.1 EBT: (a) 2030; (b) 2035; (c) 2050

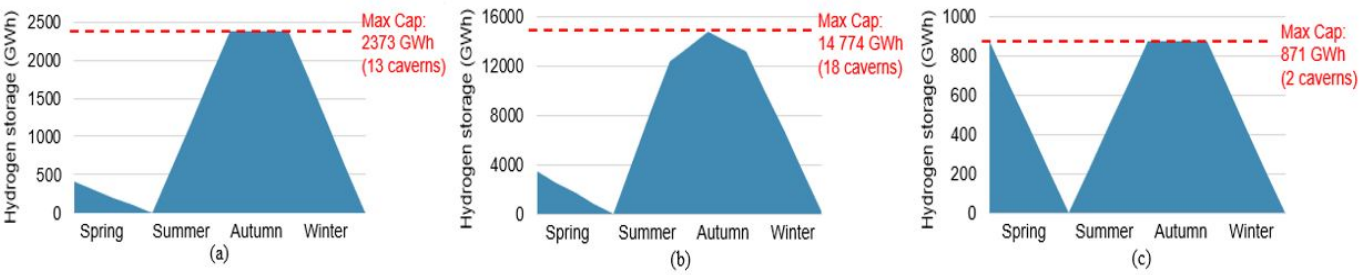


Fig. 5: Optimal hydrogen inventory in salt caverns over a typical year: (a) Cheshire underground storage in Zone 2; (b) Yorkshire underground storage in Zone 3; (c) Weald underground storage in Zone 9

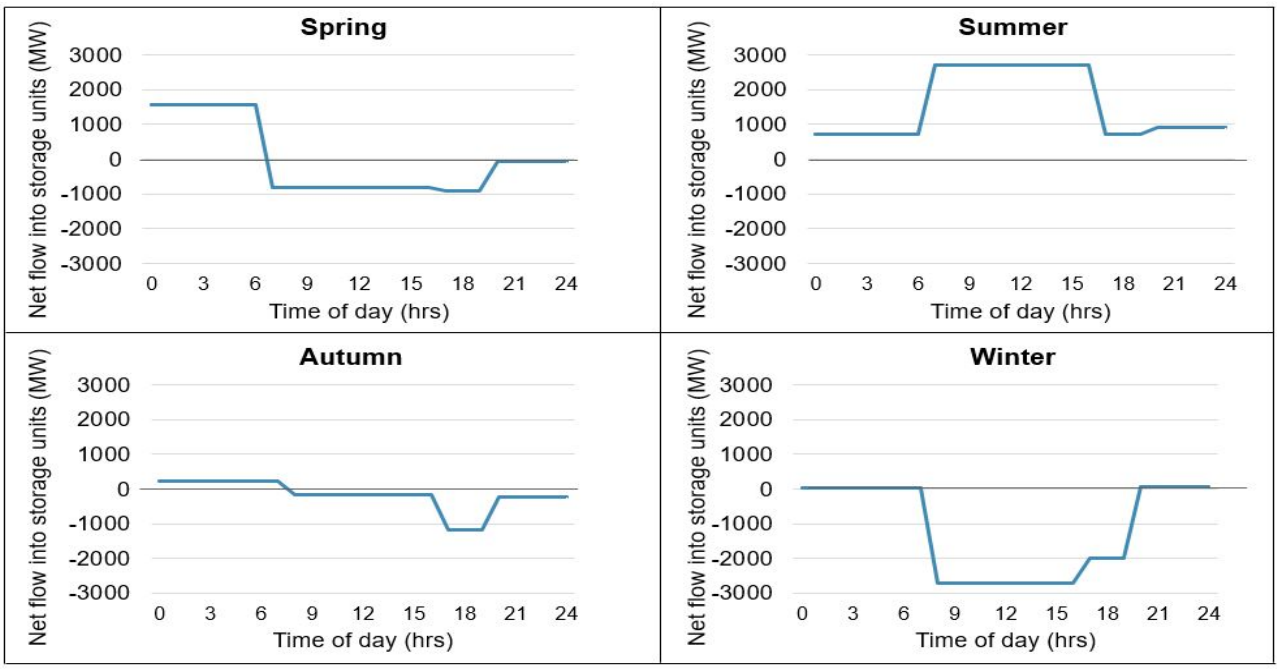


Fig. 6: Intraday storage - optimal operation of hydrogen storage inventory flows for each season

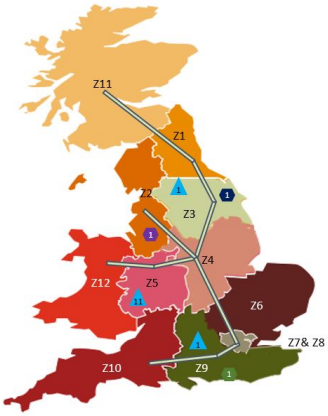


Fig. 7: Optimised spatial network for a 20mol% hydrogen gas blend at 2050

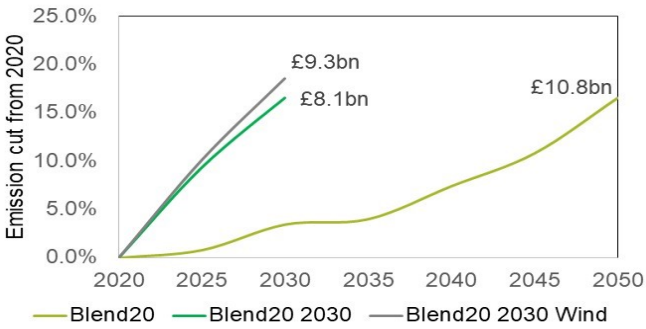


Fig. 8: Methane and Hydrogen blends cutting emission over time. Each scenario incurs a different TDC

build 11 SMRCCS units, concentrated at the centre of the country (Z5), and will invest in the pipeline network to distribute to all the zones. However, a 16% emission cut by 2050 will leave the UK far from meeting the CCC net zero GHG target. As a result of this, the model was altered to achieve this by 2030, which would result in a lower TDC of £8.1bn, as seen from Fig. (8). However, HyDeploy is currently using electrolysis to produce the hydrogen and a major assumption of the model is that CCS will be feasible before 2030. Thus, a third run which prohibited the use of SMRCCS before 2030 was executed, which resulted in a TDC of £9.3bn. This is a significant increase but it is a more realistic outcome. Overall, if HyDeploy's test are successful, a "low-regret" solution would be to scale up the 20mol% hydrogen blend with ONWTE using PEM electrolysers to achieve an 18% emission cut by 2030. This would lead to an additional cost of £38.4/tCO_2 saved. This falls cheaper than new build coal CCS on McKinsey's abatement curve [24]. It would have an LCOH of £37.4/MWh which is only a 12% increase from the nominal price.

4.3. Case 2 - Wind Outage

In order to decarbonise more than 80% of the heat demand, wind turbines with electrolysers must play a key role in the hydrogen network. However, the main concern faced by wind energy is the unpredictability in weather. Thus, an extreme case of an extended period with no wind was modelled. This period was chosen to be in the summer time slices (when wind speeds tend to be lower) of 2045 (when ONWTE already has a high adoption in the hydrogen network). The model was augmented by setting $PRDFACTOR_{ph}$ as zero for $p =$ ONWTE, OFFWTE during the summer time slices.

Fig. (9) displays the impact this would have on hydrogen inventory levels. Unlike before, storage does not build up during the summer since there is no production from wind turbines. The loss is offset from 2 factors - accumulation of storage in the spring season beforehand and increased production from SMRCCS during the summer itself. After the extended period with no wind, ONWTE would increase production (close to its maximum capacity) to bring hydrogen inventory back to that of the levels experienced in base case. This allows the EBT to still be met. Therefore, the key interpretation to be drawn from this case is that hydrogen storage is not only essential to meet demand fluctuations, but also to react to potential inconsistencies in wind speeds. This can also provide certainty in commercial contracts for utility providers, local authorities and investors as a known amount of hydogen supply can be guaranteed through storage.

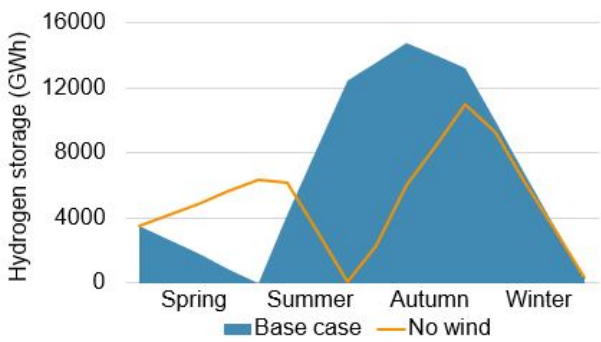


Fig. 9: Optimal hydrogen inventory in Yorkshire salt caverns (Zone 3) for base case and case with extended period of no wind during the summer of the 2045 planning period interval

4.4. Case 3 - Grid Integration

Since ONWTE would not produce up to its installed capacity at periods of lower heating demand, this case study examines the option of allowing a fraction of the wind turbine's installed capacity to be used in selling electricity to the grid. This can further incentivise the uptake of ONWTE, which was a challenge in the base case due to its higher costs relative to SMRCCS. Fig. (10) displays a significant reduction in TDC for deep decarbonisation scenarios (when EBT is less than 0.2). Naturally, the percentage of the wind turbine's installed capacity sold into the electricity grid would be limited by the demand, but this can be considered in advance when forming the generation contracts. As a result, Fig. (11) shows that when more wind capacity is sold into the grid, the earlier ONWTE appears in the transition to the hydrogen network (2025). For a 7% wind capacity sold back into the grid, it would cost £40.4/tCO_2 with LCOH £50.8/MWh for complete decarbonisation (EBT=0), much lower than the base case (£56.5/tCO_2 and £58.1/MWh).

A top down policy could encourage the construction of ONWTE units. This can be done by registering the onshore wind turbines on the Balancing Mechanism and contracting a certain amount of energy to be supplied to the electricity grid. In addition to this, there is currently 1.5TWh/year of wind energy which is curtailed [25]. This can be utilised as cheaper electricity from existing wind turbine to feed PEM electrolysers. Such combination and smart control can maximise the synergies between the heat and electricity grid, without adding large loads on to the electric grid.

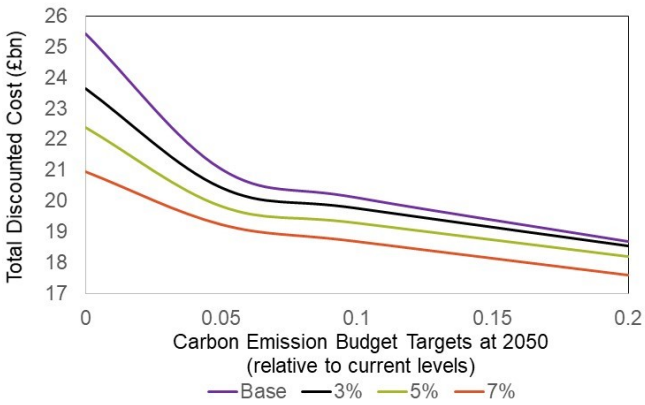


Fig. 10: Pareto solutions for deep decarbonisation scenarios, with different fractions of wind capacity sold to electricity grid

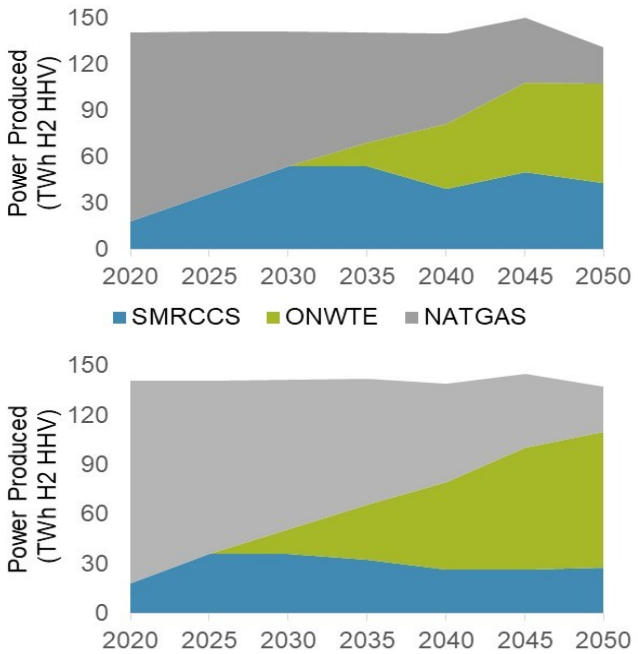


Fig. 11: Top: 3% of wind capacity sold, Bottom: 5% of wind capacity sold

4.5. *Physical Constraints*

In the base case and case studies generated by the HISTO model, SMRCCS is widely adopted. However, storage of CO_2 has not been achieved at a large scale yet. Currently, this is the largest underlying assumption which makes most of the transition scenarios feasible. Even if capture technologies, especially MEA absorption, are assumed to be feasible at large scale, the HISTO model does not consider the cost of transporting the captured CO_2 to underground formations and depleted oil and gas fields for long-term storage. By using a coarse grid, the HISTO model does not consider the associated appliance and transmission overhaul. In addition to this, there will be pressure differentials between the different levels of transmission which require compressors and expanders. This will create a further differentiation in feasibility especially when considering Case 1, since the 20mol% blend does not add such associated costs. The final output from the HISTO model relies on salt caverns for hydrogen storage. Whilst there are salt caverns around the world containing hydrogen, further safety testing is required on the UK salt caverns to ensure that capacities stated are all considered safe to provide similar load shifting flexibility as natural gas.

5. Conclusions

In this paper, the HISTO model was developed to explore a gradual transition from using natural gas to hydrogen for the provision of heat in the UK. The HISTO model comprises of a mixed-integer linear programming algorithm that provides the most cost-efficient method for designing a hydrogen infrastructure supply chain network to achieve a certain emission budget target at 2050. The model provides an optimal solution regarding the location and number of new hydrogen production, storage and transmission facilities installed across 5-year planning period intervals. In addition, the operation of the hydrogen network is resolved on an interseasonal and intraday basis.

Valuable insights can be attained from this study. As the heating demand becomes increasingly carbon constrained, there is a linear increase in TDC until EBT = 0.2. Below this target, there is an exponential increase in costs as ONWTE is adopted. The effectiveness of projects currently being piloted can be evaluated with the model. In particular, HyDeploy is a ”low-regret” solution which can be seen as a near-term strategy to remove at least 16% of the current annual CO_2 emissions within a decade. This will require a LCOH of £37.4/MWh which is a 12% increase from the current nominal price. However for deeper decarbonisation initiatives, ONWTE plays a key role in the hydrogen network. In these circumstances, underground salt caverns for hydrogen storage are crucial for balancing heat demand fluctuations at different seasons and times of the day, and alleviating the risk of wind speed unpredictability. Nevertheless, the cases have illustrated cost implications of using ONWTE. Thus, a suitable option would be to offset the costs by integrating a minor part of the heat grid with the electricity grid without overloading the latter. This can be done through careful planning and tools used by the National Grid.

The approach and results discussed in this study can benefit various stakeholders such as: (i) local planning authorities as insights into best locations and years for hydrogen technology deployment are provided; (ii) policy makers as live

pilot projects and integration of heat and electric grids are evaluated; (iii) investors as they can filter out promising hydrogen technologies; (iv) research institutions that can accelerate experiments and trials on aspects like ensuring underground salt caverns are viable for hydrogen storage on a large scale.

If physical constraints as highlighted in the paper are overcome, a hydrogen-for-heat transition would be a viable course of action into a low-carbon future.

6. Outlook

To enhance the detail of the HISTO model, CCS pipelines can be added as per Moreno's model [11]. The associated costs will alter the Pareto frontier and reduce uptake of SMRCCS. An additional case which can be further examined is a peak demand stress test, similar to the "Beast from the East" peak winter demand in 2018. For different residential and industrial properties, various potential strategies can be adopted as discussed in the Literature Review. By adding HISTO to a wider model that includes heat networks, improvements in building efficiency, heat pumps and hybrid pumps, a more detailed spatio-temporal transition analysis can be carried out.

Transportation of hydrogen is a problem due to its low density which makes it easy to leak. The HISTO model also showed that building pressurised vessels for hydrogen is expensive and will avoid doing this. As a result, the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Australia have launched a pilot project to convert hydrogen to ammonia which enables the use of existing infrastructure as well as liquid storage of ammonia [26] . If feasible, it can be added as an extension to the HISTO model.

In terms of model enhancements, using the demand and EBT parameters as a dual variable can provide the hydrogen price and carbon tax. This is complicated to attain for MILP problems but approximations for dual prices can be made.

7. Supplementary Material

Supplementary material associated with this paper can be found by clicking here.

References

[1] S. D. Watson, K. J. Lomas, and R. A. Buswell, "Decarbonising domestic heating: What is the peak GB demand?" *Energy Policy*, vol. 126, no. December 2018, pp. 533–544, 2019. [Online]. Available: https://doi.org/10.1016/j.enpol.2018.11.001

[2] BEIS, "UK Greenhouse Gas Emissions-Provisional Figures 2017," *Department for Business, Industrial and Environmental Strategy*, no. March, p. 12, 2018. [Online]. Available: https://www.gov.uk/government/collections/final-

[3] CCC, "Next Steps for UK Heat Policy," no. October, pp. 1–104, 2016. [Online]. Available: https://www.theccc.org.uk/publication/next-steps-for-uk-heat-policy/

[4] M. Hemsley, "Cleaning up the UK's heating systems: new insights on low-carbon heat," p. 1, 2018. [Online]. Available: https://www.theccc.org.uk/2018/09/10/cleaning-up-the-uks-heating-systems-new-insights-on-low-carbon-heat/

[5] R. Loulou, G. Goldstein, and K. Noble, "Documentation for the MARKAL Family of Models," *International Energy Agency*, no. October, pp. 1–389, 2004.

[6] Ofgem, "The Decarbonisation of Heat," *Ofgem's Future Insights Series*, pp. 3–4, 2016. [Online]. Available: https://www.ofgem.gov.uk/system/files/docs/2016/11/ofgem_future_insights_programme_the_decarbonisation_of_heat.pdf

[7] Element Energy, "BEIS - Hydrogen (H2) supply chain evidence base," no. November, 2018.

[8] D. Sadler, A. Cargill, M. Crowther, A. Rennie, J. Watt, S. Burton, and M. Haines, "H21 Leeds City Gate Report," pp. 1–382, 2016. [Online]. Available: https://www.northerngasnetworks.co.uk/wp-content/uploads/2017/04/H21-Report-Interactive-PDF-July-2016.compressed.pdf

[9] T. Isaac, "HyDeploy: The UK's First Hydrogen Blending Deployment Project," *Clean Energy*, vol. 3, no. 2, pp. 114–125, 2019.

[10] A. Almansoori and N. Shah, "Design and operation of a future hydrogen supply chain: Snapshot model," *Chemical Engineering Research and Design*, vol. 84, no. 6 A, pp. 423–438, 2006.

[11] M. Moreno-Benito, P. Agnolucci, and L. G. Papageorgiou, "Towards a sustainable hydrogen economy: Optimisation-based framework for hydrogen infrastructure development," *Computers and Chemical Engineering*, vol. 102, pp. 110–127, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.compchemeng.2016.08.005

[12] M. Dayhim, M. A. Jafari, and M. Mazurek, "Planning sustainable hydrogen supply chain infrastructure with uncertain demand," *International Journal of Hydrogen Energy*, vol. 39, no. 13, pp. 6789–6801, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.ijhydene.2014.02.132

[13] J. Kim and I. Moon, "Strategic design of hydrogen infrastructure considering cost and safety using multiobjective optimization," *International Journal of Hydrogen Energy*, vol. 33, no. 21, pp. 5887–5896, 2008. [Online]. Available: http://dx.doi.org/10.1016/j.ijhydene.2008.07.028

[14] N. Sabio, A. Kostin, G. Guillén-Gosálbez, and L. Jiménez, "Holistic minimization of the life cycle environmental impact of hydrogen infrastructures using multi-objective optimization and principal component analysis," *International Journal of Hydrogen Energy*, vol. 37, no. 6, pp. 5385–5405, 2012.

[15] A. Lahnaoui, C. Wulf, H. Heinrichs, and D. Dalmazzone, "Optimizing hydrogen transportation system for mobility by minimizing the cost of transportation via compressed gas truck in North Rhine-Westphalia," *Applied Energy*, vol. 223, no. May, pp. 317–328, 2018. [Online]. Available: https://doi.org/10.1016/j.apenergy.2018.03.099

[16] M. Reuß, T. Grube, M. Robinius, and D. Stolten, "A hydrogen supply chain with spatial resolution: Comparative analysis of infrastructure technologies in Germany," *Applied Energy*, vol. 247, no. December 2018, pp. 438–453, 2019. [Online]. Available: https://doi.org/10.1016/j.apenergy.2019.04.064

[17] S. Samsatli and N. J. Samsatli, "The role of renewable hydrogen and inter-seasonal storage in decarbonising heat – Comprehensive optimisation of future renewable energy value chains," *Applied Energy*, vol. 233-234, no. September 2018, pp. 854–893, 2019. [Online]. Available: https://doi.org/10.1016/j.apenergy.2018.09.159

[18] BEIS, "Regional, local authority gas consumption statistics," 2018. [Online]. Available: https://www.gov.uk/government/statistical-data-sets/gas-sales-and-numbers-of-customers-by-region-and-local-authority

[19] Pure Energy Centre, "Launch Of The Greenest Hydrogen Boiler In The World," 2012. [Online]. Available: https://pureenergycentre.com/hydrogen-boiler-pure-energy-centre-product-launch/

[20] G. Sinden, "Characteristics of the UK wind resource: Long-term patterns and relationship to electricity demand," *Energy Policy*, vol. 35, no. 1, pp. 112–127, 2007.

[21] S. Samsatli, I. Staffell, and N. J. Samsatli, "Optimal design and operation of integrated wind-hydrogen-electricity networks for decarbonising the domestic transport sector in Great Britain," *International Journal of Hydrogen Energy*, vol. 41, no. 1, pp. 447–475, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ijhydene.2015.10.032

[22] BEIS, "Clean Growth - Transforming Heating - Overview of Current Evidence," no. December, p. 13, 2018. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment{_}data/file/766109/decarbonising-heating.pdf

[23] S. Samsatli and N. J. Samsatli, "A multi-objective MILP model for the design and operation of future integrated multi-vector energy networks capturing detailed spatio-temporal dependencies," *Applied Energy*, vol. 220, no. July 2017, pp. 893–920, 2018. [Online]. Available: https://doi.org/10.1016/j.apenergy.2017.09.055

[24] K. Gillingham and J. H. Stock, "The cost of reducing greenhouse gas emissions," *Journal of Economic Perspectives*, vol. 32, pp. 53–72, 2018.

[25] B. Coyne, "3% of Britain's wind power 'wasted' in 2017," 2018. [Online]. Available: https://theenergyst.com/3-uk-wind-power-wasted-2017/

[26] M. Gillespie, "Hyper for hydrogen: our world first for carbon-free fuel," 2018. [Online]. Available: https://blog.csiro.au/hyper-for-hydrogen-our-world-first-carbon-free-fuel/

Mechanistic Modelling of RNA Vaccine Production for Quality by Design

Damien van de Berg and Carl Fredrik Behmer

Department of Chemical Engineering, Imperial College London

Abstract: The Future Vaccine Manufacturing Research Hub is working on enabling rapid distribution of low-cost vaccines in response to the outbreak of epidemics, especially in low-income countries. Designing a robust and modular platform for the production of RNA vaccines has the potential to achieve this target by reducing development time, equipment and logistics cost. By implementing a Quality by Design framework, the efficacy and safety of the vaccine product, as well as the robustness of the process can be ensured. This study identifies the critical qulity attributes (CQA) and applies a novel framwork to identify the corresponding crtical process parameters (CPP). A mechanistic model is developed that links the CQAs of the transcription reactor to the CPPs. This allows for the creation of the first deterministic, as well as probabilistic design space of an RNA vaccine transcription unit using mechanistic modelling. The mechanistic model is successful in showcasing cause-effect relationships and strengthening process understanding in combination with experimental design of experiments. However, the model is unable to predict process behavior in the unexplored regions with potentially different mechanisms. Nevertheless, the model is a viable alternative to statistical models if faced with limiting data. The model lays down the foundation for process development, optimisation and control. Furthermore, this work investigates how the model can be used in the design of a practical continuous-flow microreactor.

Keywords: RNA **V**accines, In Vitro Transcription, Mechanistic Modeling, Quality by Design, Design Space

1. Introduction

The approval and effective distribution of vaccines can prevent the outbreak of epidemics [1]. However, the high cost of vaccine development and production can often prevent their distribution in low-income countries [2], which motivates research into methods of producing safe and effective vaccines at a low cost. RNA vaccines are an exciting area of research because they are inherently safe and could potentially reduce the development time from years to weeks [3]. RNA *in vitro* transcription constitutes one process unit within a preliminary RNA vaccine production process [2]. In the transcription reactor, nucleotide triphosphates (NTP^{4-}) are transcribed into RNA in an enzymatic reaction including T7RNAP and the DNA template in a complex HEPES buffer solution including over 15 components, of which the most important are Mg^{2+}, NTP^{4-}, spermidine and DTT. Finally, RNA vaccine manufacturing processes are inherently flexible because they can be applied to combat different viruses by only changing the DNA sequence [4].

Currently, the quality of pharmaceutical products is mostly assured through Quality by Testing (QbT) [4]. QbT follows a sequence of steps where the raw material, process, and product are tested separately. If all quality controls are passed, the product can be released to the market. Otherwise, the control must be repeated from scratch. One of the major shortcomings of quality by testing is that it generates no insight into what went wrong in the production process [4], creating the risk of the errors being repeated. The introduction of Quality by Design (QbD) as a replacement to QbT happened initially in 2004 [5]. Successfully implementing QbD into the production process reduces both the required quality testing and the risk of a product failing such testing. The QbD framework consists of six distinct steps [4]:

1. Define the Quality Target Product Profile (QTPP) based on the patients' needs.

2. Identify the Critical Quality Attributes (CQA) and Critical Process Parameters (CPP).

3. Establish relationships between the CQAs and CPPs

4. Define the design space that maps the range of operating conditions that ensures the QTPP

5. Develop a control system for the process

6. Continuously monitor the process and feed the process parameters to the control system.

By knowing the feasible design space of the process, the quality of the product is guaranteed provided that the operating point does not exit this space at any instance of the production process.

Finding the design space experimentally can be a both time-consuming and expensive task, especially if there are several potential CPPs. Multivariate statistical modelling is the most common computational tool to correlate CQAs and CPPs. This type of model has high predictive power, and has the advantage of quantifying uncertainty for risk management. However, it may fail to illustrate causality [6], and thus does not provide a deeper understanding of the process, which constitutes one of the pillars of QbD in addition to risk management. The development of a computational model based on a mechanistic understanding enables for a quick yet robust analysis of all potential CPPs. A mechanistic model is superior to a statistical model due to its clear showcasing of cause and effect relationships through physical terms. The model can be tuned and validated by performing fewer experiments. Once a model with proven predictive power is developed, the design space can be generated and the optimised normal operating range within the design space can be found. Furthermore, by assigning normally distributed uncertainties to the process parameters, a probabilistic design space can be generated, which is useful for the mapping of the normal operating range.

2. Background

The Future Vaccine Manufacturing Research Hub (FVMRH) is a multi-institutional research organisation exploring the development of cheap, reliable and flexible novel manufacturing platforms with the aim to quickly respond to potential outbreaks [7]. One of many platforms explored by the FVMRH is RNA vaccines production. Multiple groups collaborate to optimise the transcription and tangential flow filtrations steps, which is only a smaller part of the overall effort. The most interesting research contribution for this report comes from Professor Shattock's group performing Design of Experiments (DoE) for the *in vitro* RNA vaccine transcription reaction[1].

Attempts to create computational models following the QbD framework for bioprocess control have been made. Jedrzejewski, del Val, Polizzi, and Kontoravdi performed experiments and developed a mathematical model for the control of protein glycosylation [8]. Akama, Yamaura, and Kigawa developed a biophysical model for RNA transcription and magnesium precipitation but their model did not follow the principles of Quality by Design, nor did they incorporate RNA degradation [9]. Haas et al. published a case study on the implementation of QbD in the development of vaccines [10]. However, creating a design space for the *in vitro* transcription of RNA using mechanistic modelling has never been done before. This sets the scene for an exploration of using mechanistic modelling in QbD for the RNA vaccine production process, more specifically, in creating a design space for the *in vitro* RNA vaccine transcription. This work can then be used in the design of an idealised continuous plug flow reactor (PFR), with the aspiration to design a modular production system, able to quickly produce low-cost vaccines in low-income countries.

3. Methodology

3.1. Identification of the CQAs

In accordance with the QbD framework, the initial step when developing the model is to identify the Critical Quality Attributes (CQAs) [5]. An exhaustive list of quality attributes (QAs) for the specific process was derived from the QTPP provided by Professor Shattock, see **Table 1**. In this table the QAs were ranked by adapting an existing risk-assessment based framework developed by GSK [11][4]. Each QA is given two semi quantitative scores, an impact score (2 for low impact, 5 for moderate impact, or 25 for high impact), and an uncertainty score (ranging from 1 for low uncertainty up to 5 for high uncertainty). These two scores are then factored together, and the product is the severity score.

If the severity score is 10 or below, then the QA is not critical and will thus not be considered in the model. If the severity score is 16 or 24 (resulting from an impact score of 8 and an uncertainty score of 2 or 3) then the QA can be, but is not necessarily, critical and the nature of the clinical information available will be considered to determine the criticality of the QA. If the severity is 25 or above, then it is the uncertainty score that determines the criticality of the QA. If the uncertainty score is between 1 and 3, then a severity score of 25, 30 or 75 is obtained, and the QA is a CQA, and will be considered in the model. If the uncertainty is 4 or 5, then the severity score is 32, 40, 100 or 125, then the QA is a potential CQA, and will be considered in the model, but not given the same weight as an absolute CQA [4].

The production process is divided up into two sections: the upstream process where the RNA is

[1] Data from Professor Shattock's group that is not yet published will be referred to throughout the article

produced and the downstream process where the RNA is purified and formulated. The key unit operation in the upstream process is the *in vitro* transcription (IVT) reaction and the key unit operation in the downstream process is the tangential flow filtration (TFF). These two key unit operations would initially be modelled separately, starting with IVT, then continuing with the TFF and finally generating the whole RNA production process model. This way, the current focus is on developing the QbD model for the IVT reaction and for this criticality of QAs presented in **Table 1** has been assessed. Examples of QAs not considered in the transcription reaction is the product pH (as this can be adjusted later) and the amount of precipitate (as this is removed during the filtration). Bacterial endotoxins and bioburden are not considered either, based on the assumption that the production process is operated under sterile conditions. Left as CQAs of the transcription reactions are: Yield, RNA sequence integrity (length and stability of the molecule), and sequence identity (are the nucleotides added in the right order).

Table 1. Criticality assessment for quality attributes (QAs) relevant for the *in vitro* transcription reaction in the upstream bioprocess for RNA vaccine production.

QAs	**Severity Score - Safety** [$]	**Severity Score – Efficacy** [$]	**Max Severity** [$]	**Classification** [€]
Amount of precipitate	8	32	32	pCQA
RNA sequence integrity	32	75	75	CQA
RNA sequence identity	32	75	75	CQA
Content/concentration (RNA)	6	24	24	CQA
Residual host cell proteins (E. Coli)	6	6	6	QA
Residual host cell DNA (E. Coli)	24	6	24	QA
Residual template DNA	24	6	24	QA
Bacterial endotoxins	50	8	50	CQA
Bioburden	50	32	50	CQA
pH post filtration	6	75	75	CQA
Salt concentration	6	24	24	QA

[$] The severity score was obtained by multiplying the impact score and the uncertainty score.

[€] The quality attributes are classified into: critical quality attributes (CQAs, in red), potential critical quality attributes (pCQAs, in yellow), and non-critical quality attributes (QAs, in green)

3.2 Identification of the CPPs

The next step after identifying the CQAs is to identify the process parameters (PPs) of the transcription reaction, assess their criticality, and tie CPPs to the CQAs. A novel groundwork to allow for such mapping is presented, where each PP is assigned a number ranging from 0 to 3 and a sign label of plus, plus-minus, or minus. The sign label represents the expected derivative of the CQA performance as a function of the respective PPs.

Plus indicates that the derivative is always greater than or equal to zero (could be both constant or tend towards zero). Plus-minus indicates that there is an expected optimum value for the PP, and deviations from this expected value will worsen the performance of the CQA. Minus implies that the derivative of the CQA is strictly negative in terms of the PP. The number represents how sensitive the CQA is expected to be to changes in the PP. PPs that are assigned values of 2 or 3 are labelled CPPs and are considered for the model.

3.3 Linking CQAs to CPPs

3.3.1. Model assumptions

Three CQAs are identified, namely RNA yield, RNA sequence integrity and sequence identity. Due to limited available literature on the mechanism of incorrect transcription, sequence identity is omitted from the model. The two remaining CQAs, yield and sequence integrity, are incorporated into one effective CQA, the effective RNA concentration.

Since the aim of this investigation is to elucidate the concept of using mathematical models in QbD to create a design space, the model is simplified for the sake of clarity and visualisation. Hence, temperature is assumed to be constant, the DNA template is assumed to have the right sequence, m7G cap concentration is neglected, and the four NTP^{4-} concentrations are assumed to be equimolar.

A key objective of this project is to move away from statistical models derived from experimental data and instead establish mathematical relations between the CPPs and the CQAs, based on a mechanistic understanding of the process. DTT and spermidine concentrations, even though determined experimentally to be critical to the operation, are not included as variables in the model. Intestad they are assumed to be at their experimental optimum throughout the whole reaction, since there is a lack of mechanistic understanding about their roles in *in-vitro* RNA transcription.

Hence, the change in effective RNA yield is linked to reaction time, total Mg^{2+} concentration, total NTP^{4-} concentration and T7RNAP concentration. The total proton concentration appears in the model through its dependency on the change in effective RNA yield.

3.3.2. Model structure

The model consists of a system of Ordinary differential equations (ODEs), describing the change in total RNA, PPi^{4-}, H^+, and NTP^{4-} concentration through a mechanistic transcription and for RNA also a mechanistic degradation term. The initial transcription term is adopted from Akama et al. and follows Michaelis-Menten kinetics with Mg^{2+} and $MgNTP^{2-}$ as substrates [9]. In later iterations, substrate inhibition terms are included in the denominator of the transcription term. Inspired by Li, the degradation term is described as a power law including RNA and proton, hydroxyl or Mg^{2+} concentration [10]. To avoid overfitting, the reaction orders are assumed to be first order with respect to RNA, as well as either H^+, OH^- or Mg^{2+} concentration. The transcription term represents the CQA of RNA yield while the degradation term represents sequence integrity.

$$\frac{d[RNA_{effective}]}{dt} = V_{trans} - V_{deg} \quad (3.3.1)$$

$$\frac{d[C_i]}{dt} = K_i\, V_{trans},\ \mathrm{i} \in \{[\mathrm{PPi}^{4-}], [\mathrm{H}^+], [\mathrm{NTP}^{4-}]\} \quad (3.3.2)$$

$$V_{Trans} = \frac{k_{app}[T7RNAP][MgNTP^{2-}][Mg^{2+}]}{1 + K_1[MgNTP^{2-}] + K_2[Mg^{2+}] + K_3[Mg^{2+}]^2 + K_4[MgNTP^{2-}]^2 + K_5[Mg^{2+}][MgNTP^{2-}]^2}$$

(3.3.3)

$$V_{deg} = \left(k_{ac}[H^+]^{n_{ac}} + k_{ba}[OH^-]^{n_{ba}} + k_{Mg}[Mg^{2+}]^{n_{Mg2+}}\right)[RNA] \quad (3.3.4)$$

In contrast to Akama, the reaction considered in this report is not substrate-limited within the considered time frame. Therefore, no artificial T7RNAP degradation term is included. Reaction times above six hours are not considered for economical and logistical reasons, and to avoid excessive RNA degradation.

Precipitation is omitted from the model since it only has a minor effect on the total concentration of Mg^{2+} and on the transcription kinetics, even though it is a crucial physical phenomenon that thermodynamically drives the transcription reaction[9].

3.3.3. Numerical methods

Following Akama's method at each time step, the free solution concentrations are computed through solving a system of non-linear equations given by mass balances and equilibrium constraints[9]. These free solution concentrations then appear in the ODE expressions determining the change in total concentration. This way, a complex system of high-index differential-algebraic equations is split up into two different problems: Solving a system of initial value ODEs and solving a system of non-linear equations at each time step. The ODEs are solved explicitly through implementing the 4th order Runge-Kutta solution method. The non-linear equations are solved using the built-in *scipy.optimize.fsolve()* function in Python. The computation of free solution concentrations depends on the initial guesses for the ODEs. To avoid the convergence towards unphysical solutions that would blow up the ODEs, an algorithm was developed to explore different initial guesses and update the initial guess to the iteration once a physical solution is found.

3.3.4. Parameter estimation

Validation of a mechanistic model usually follows these steps: An initial guess is made regarding the form of the mechanistic terms and the parameters are then estimated to fit experimental results. The model is then validated against additional experimental data.

Experimental transcription data was made available from Professor Shattock's research group exploring the design of experiments in RNA transcription. The data sets include experimental results over a wide range of process conditions. In the end, 45 sets are used at pre-determined optimal spermidine, DTT and monovalent cation concentration, to investigate the dependence of effective RNA yield on Mg^{2+} and T7RNAP enzyme concentration at 40 mM NTP^{4-}. The resulting model is then qualitatively validated against the experimental RNA yield of six sets of conditions: NTP^{4-} concentrations of 20, 40 and 80 mM at Mg^{2+} concentrations of 75 and 140 mM. This method of validation is chosen due to limited availability of data. If more data, especially over more NTP^{4-} ranges had been available, data could

have been fitted over one proportion, say 95% of the data, and then cross-validated against the other 5%.

The parameters k_{app}, K_1, K_2, K_3, K_4, K_5 , k_{ac}, k_{ba}, k_{Mg2+} are then estimated using the scipy.*curve_fitting()* function in Python through least-squares error minimisation. It is noteworthy to mention that *curve_fitting()* is only a local solver and thus is highly dependent on the initial guess provided.

3.4. Mapping the design space

3.4.1. Deterministic 3D-design space

A deterministic model can then be produced by adopting the estimated optimum parameters from the data fitting section. The creation of the deterministic model starts with discretising the 3D space into a 3D grid with Mg^{2+} and NTP^{4-} concentrations on the x- and y-axis and T7RNAP concentration on the z-axis. Each grid point represents a set of initial concentrations of Mg^{2+}, NTP^{4-} and T7RNAP, and is only plotted if it meets a CQA target of 1.5g/L effective RNA yield after 6 hours. The target CQA is set arbitrarily and can be revised after creating the design space for the other RNA vaccines production units. The design space is colour-coded indicating the performance of the reaction at that point – bright yellow indicates that the 1.5g/L target is just met, whilst dark green corresponds to RNA yields of up to 6 g/L.

3.4.2. Probabilistic 2D-design space

In a deterministic model, the implication is that the model is accurate. However, creating a probabilistic design space instead is a way of accounting for parameter estimation, experimental uncertainty and various unanticipated uncertainties such as process fluctuations. To this end, T7RNAP concentration is fixed at 1.0 e-7 M for all sequential purposes to alleviate computational cost.

The Mg^{2+}/NTP^{4-} space is discretised and the RNA yield after a six hour reaction time is determined in 50 simulations over the whole grid. For each simulation, the model parameters are sampled from a normal distribution with their respective estimated parameter optimum as the mean and a 20% standard deviation. Instead of adapting the standard deviations given by the curve fitting for the three different parameters, the maximum of which is 12% for k_{Mg2+}, 20% is chosen to account for the currently unquantifiable uncertainties such as process fluctuations.

The design space is again colour-coded: Grid points that have a zero probability of reaching the required CQAs are not included. Grid points that have a very slight chance of meeting the CQA are dark-red. The greener the region, the more certain a certain set of initial conditions is to always reach the CQA.

3.5 Design of a PFR RNA vaccine transcription microreactor

Following the creation of the design space, a potential RNA vaccine transcription reactor can be modelled. **Eq. (3.3.1)** and **Eq. (3.3.2)** from the current batch reactor model can easily be applied to an idealized PFR setup by replacing the reaction time by residence time.

$$\frac{d[RNA_{effective}]}{d\tau} = V_{trans} - V_{deg} \quad (3.5.1)$$

$$\frac{d[C_i]}{d\tau} = K_i\, V_{trans}, \mathrm{i} \in \{[\mathrm{PPi}^{4-}], [\mathrm{H}^+], [\mathrm{NTP}^{4-}]\} \quad (3.5.2)$$

$$Residence\ time: \tau = \frac{Reactor\ volume}{Flow\ Rate} \quad (3.5.3)$$

For instance, from the 2D-probabilistic design space in **Figure 4**, initial conditions are chosen around the experimental optimum at: 1e-7M T7RNAP, 40 mM NTP^{4-} and 85 mM Mg^{2+}

Starting from a certain reactor geometry, a flowrate can finally be found that gives a six-hour residence time. For this specific application, the following reactor dimensions are chosen from Lawrence's PhD thesis [12]: 55 sections in a "zig-zag" arrangement of cross-section 1mm*0.5mm and length 60 mm (total length 3,300 mm) for a total volume of 1.65 mL.

4. Results and discussion.

4.1. Identification of CPPs

The identified CPPs are presented in **Table 2.** A detailed reasoning regarding the criticality of each PP is presented in **Appendix A.1**

Table 2. The qualitative assigned relation between the process parameters and the CQAs. Those PPs that had an impact rating of 3 were labelled critical. A rating consisting of a number and sign(s) is provided for each CQA-PP pair in order to assess the criticality of the PP. The number (i.e. 0, 1,3) indicates the magnitude of the impact from 0 representing low to 3 representing high. The sign(s) indicates the shape of the graph when the CQA is plotted in function of the (C)PP; "+" indicating a positive slope, "-" indicating a negative slope and "±" representing a peak behaviour consisting of an increase up to a peak value followed by a decrease. A detailed reasoning regarding the classification of each PP is presented in **Appendix A.1**

PPs	RNA sequence integrity	RNA sequence identity	RNA Yield	Classification
Temperature	-2	-1	±3	CPP
Pressure	±1	±1	±1	PP
Mixing	±1	±1	±1	PP
Reactor dimensions	0	0	0	PP
Reaction time	±2	0	3	CPP
pH in transcription reactor	±2	±2	±3	CPP
DNA template sequence	-2	-3	0	CPP
DNA template conc.	0	0	+2	CPP
T7RNAP conc.	±1	±1	+3	CPP
m7G cap conc.	0	±3	0	CPP
Mg^{2+} conc.	0	±2	±3	CPP
DTT conc.	±2	±2	±1	CPP
Spermidine conc.	±2	±2	±3	CPP
GTP conc.	+1	±3	±3	CPP
Total NTP^{4-} conc.	+2	±2	±3	CPP
Ratio of NTP^{4-}s	+1	±3	±3	CPP

4.2.Parameter estimation

Through trial and error, it is found that k_{app} and k_{Mg2+}, are most important in fitting the curve. In fact, the only parameters that have any major impact are k_{app}, K_5, and k_{Mg2+} with values of 13,000 ± 70 $M^{-2}hr^{-1}$, 200,000 ± 300 M^{-3}, and 17 ± 2 $M^{-1}hr^{-1}$ respectively. The parameter with the highest relative uncertainty is k_{Mg2+} with a standard deviation of ~12%. The data fitting, as seen on **Figures 1 and 2** has a mean absolute error (MAE) of 0.27 g/L, which compares to the MAE in the experimental data provided of 0.15 g/L. MAE is chosen as an error metric since any method that normalizes the error gives larger errors. It should be reiterated that this parameter estimation might be highly biased by the choice of initial conditions.
From the data fitting in **Figure 1**, it can be noted that the model fits the experimental data after two hours quite poorly.

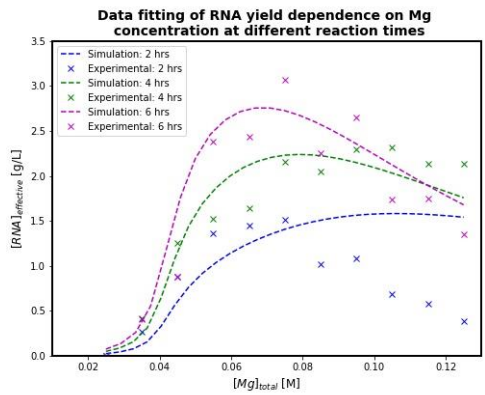


***Figure 1**: Data fiting of RNA yield dependence on Mg^{2+} concentration at 40 mM NTP^{4-} and 10^{-7} M T7RNAP at 2 (blue), 4 (green) and 6 (magenta) hours reaction time*

However, it is more important to note that the model is accurate after 4-6 hours, when the design space constraints are specified.

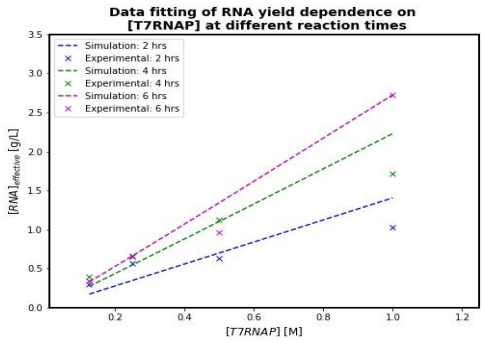


Figure 2: *Data fiting of RNA yield dependence on T7RNAP⁺ concentration at 40 mM NTP⁴⁻ and 85 mM Mg²⁺ at 2 (blue), 4 (green) and 6 (magenta) hours reaction time*

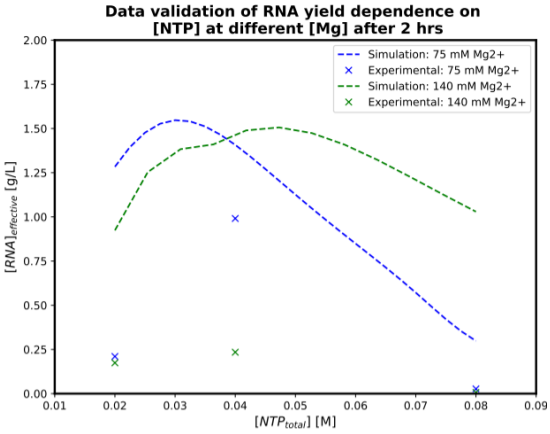


Figure 3: *Data validation plots describing RNA yield after a reaction time of 2 hours as a function of NTP^{4-} concentration at 75 mM (blue) and 140 mM (green) RNA*

Although the simulated and experimental validation values deviate significantly **(Figure 3)**, they both follow the same trend and predict an optimum NTP^{4-} concentration at around 40 mM at fixed Mg^{2+} concentration. Since no combination of the coefficients in **Eq. (3.3.3)** and **Eq. (3.3.4)** are found to even fit the general trend of the validation data, k_{app}, K_5, and k_{Mg2+} are chosen since they best describe the overall trend. However, as illustrated in **Figure 1**, the model results after 2 hours might not accurately predict whether the results are correct after six hours. This means that the validation data might not be reflective of the outcome of a reaction time of 6 hours and highlights the need for NTP^{4-} validation data after 6 hours at different Mg^{2+} concentrations.

The model describes the region around the experimental optimum conditions 80 mM Mg^{2+} and 40 mM NTP^{4-} after six hours most accurately. The uncertainty of the model increases as the NTP^{4-} concentration deviates from 40 mM. This illustrates the main drawback of mechanistic models. It is difficult to construct a model which accurately predicts the behaviour outside of the region in which the parameters were estimated, since mechanisms tend to change over wide ranges.

4.3. Mapping the design space

4.2.1. 3D deterministic design space

There are some key results that should be highlighted from **Figure 4**. With the mentioned coordinate system, the design space is shaped like a reverse pyramid, which is greener in the centre. This indicates that there is a positive relation between T7RNAP concentration and yield, as the production target is met for a wider range of Mg^{2+} and NTP^{4-} concentrations. At constant T7RNAP, there appears to be an optimum ratio between Mg^{2+} and NTP^{4-} concentration, seen as the dark green region in the centre of Mg^{2+}/NTP^{4-} plane. The yield seems to first increase with increasing Mg^{2+} and NTP^{4-} concentration and then decrease after a global optimum is reached.

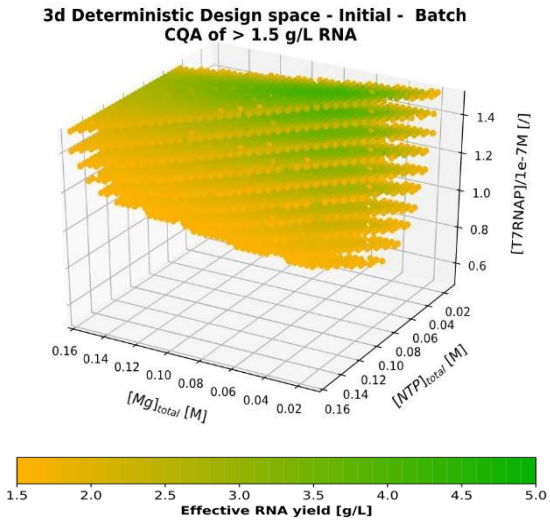


Figure 4: *3D deterministic design space mapping the effect of CPPs Mg^{2+}, NTP^{4-} and T7RNAP concentration. Every point within the design space meets a production target of 1.5 g/L RNA. The greener the region, the higher the yield*

The positive correlation between T7RNAP concentration and RNA yield is expected. The T7RNAP concentration appears as a first order reactant in the transcription reaction **(Eq. 3.3.3)** and does not contribute to the degradation of transcribed RNA. However, as T7RNAP is one of the major cost drivers of the reaction, its optimum concentration can be determined through cost optimisation.

The optimum in NTP^{4-}, at fixed Mg^{2+}concentration may appear less intuitive, but can be seen clearly in **Figure 5**. At first, increasing NTP^{4-} concentration leads to a higher concentration of $MgNTP^{2-}$, which in turn increases the reaction rate. However, if the Mg^{2+} concentration is not sufficiently high, an increase in the NTP^{4-} concentration causes a

proportionally smaller increase in the $MgNTP^{2-}$ complex concentration compared to the reduction of the Mg^{2+} concentration. Since the transcription reaction has a first order dependency on both species, the effect of increasing the NTP^{4-} concentration above the optimum is a reduction in the transcription rate.

Yet, at high Mg^{2+} concentrations, the optimum in NTP^{4-} cannot simply be explained by this relative decrease in free solution Mg^{2+} concentration. Therefore a $[Mg^{2+}][MgNTP^{2-}]^2$ term needed to be included in the model. At both high Mg^{2+} and NTP^{4-} concentrations, **Eq. (3.3.3)** reduces to an inverse dependence on $MgNTP^{2-}$, which can be interpreted as substrate inhibition. To more accurately describe the decrease in RNA yield at higher NTP^{4-}, a potential free solution NTP^{4-} inhibition term could be investigated.

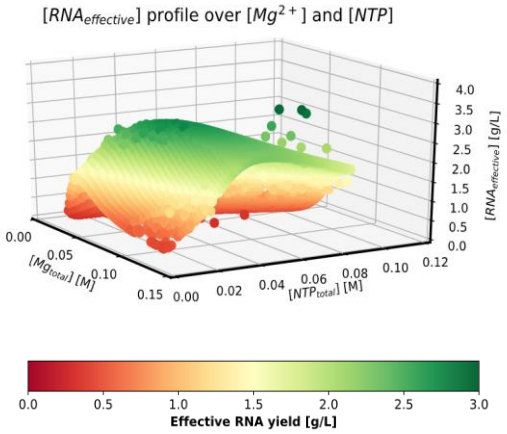


Figure 5. *3D Plot showing the effective RNA profile as a funciton of Mg^{2+} and NTP^{4-} concentration at a constant T7RNAP concentration of 1.0 e-7*

In addition to strong Mg^{2+} and $MgNTP^{2-}$ substrate inhibition, the optimum in Mg^{2+} concentration is explained by the first order dependence of the degradation on free solution Mg^{2+} concentration. The increased degradation rate counteracts the increasing transcription rate at higher total Mg^{2+} and thus higher free solution Mg^{2+} and $MgNTP^{2-}$ concentrations, which appear both in the numerator and denominator **(Eq. 3.3.3)**. At low Mg^{2+} concentrations, the numerator increases the transcription rate at increasing Mg^{2+} concentrations, while at higher Mg^{2}+ concentrations, the transcription term becomes constant and the degradation becomes dominant.

Both substrate inhibition and Mg^{2+} degradation then lead to a global optimum on a fixed T7RNAP surface.

4.2.2. 2D probabilistic design space

The 2D probabilistic design space in **Figure 6** illustrates the Mg^{2+}/NTP^{4-} optimum more effectively than the 3D deterministic version. The optimum is seen as the dark green region around 70 mM Mg^{2+} and 40 mM NTP^{4-}. It appears that the definition of the lower edge of the design space is most affected by uncertainty parameter estimation and that the upper left boundary does not change too much with parameter uncertainty. However, the cause for this might not be physical, but could be numerical. Since the cut-off seems quite abrupt, there might be a discrepancy in the computation of free solution concentrations at given initial NTP^{4-} and Mg^{2+} concentrations, or points towards a discontinuity or other shortcomings in the model. Additionally, as discussed in the parameter estimation section, the model becomes less reliable the further the NTP^{4-} concentrations deviate from 40 mM. More data is needed to fit the actual upper boundaries. The design space probably does not stretch out as far to the upper right as the model predicts. Experimental results point towards the RNA yield approaching zero at both high Mg^{2+} and NTP^{4-}concentrations.

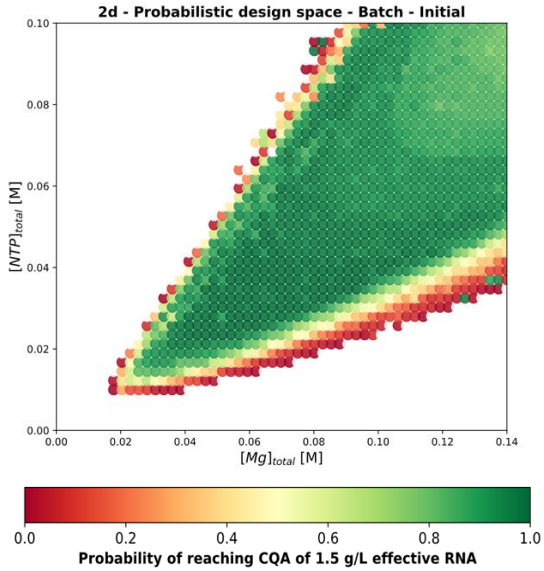


Figure 6: *2D probabilistic design space giving the probability of a grid point reaching the CQA of 1.5 g/L RNA yield given a 20% standard deviation around the estimated parameter optimum*

This type of design space is an appropriate starting point for the mapping of the normal operating range. The normal operating range is defined within the design space with a 'safety margin', to avoid that the operating point is accidentally carried outside the design space by process fluctuations. Intuitively, the darker green space could serve as a preliminary normal operating range. To achieve more accuracy, one could optimise the parameters of the equation of a rectangle within the design space to maximise the area of the normal operating range under the

condition that all interior points have a minimum probability of achieving the CQAs.

With the normal operating range defined, a control strategy can be implemented. After linking physical measurements of Mg^{2+}, NTP^{4-}, and RNA concentrations to the model, a dynamical design space can be created that takes into account current RNA concentrations and reaction time to give the CPPs that lead to the desired final RNA yield.

4.3. Design of a PFR RNA vaccine transcription microreactor

The average volumetric flowrate is computed to be 4.6 µl/min to give a residence time of 6 hrs. Assuming an ideal PFR model, the initial concentrations of 1.0 e-7 M T7RNAP, 40 mM NTP^{4-} and 85 mM Mg^{2+}, would yield 2.4 g/L at the end of the reactor.

Assuming 20 µg/dose and a continuous outlet of 2.4 g/L for all microreactors of this configuration, one single reactor would need to operate for 1.8 min to produce a single dose, or a week for 5,600 doses or 1,260 days for 1 million doses.

To put this into perspective, 'Every year more than 29 million people worldwide receive a post-bite vaccination [13]. The particularity about microreactors is their potential for scale-out. 173 microreactors can operate in parallel to produce 50 million doses of vaccines/year. Additionally, vaccines can theoretically be produced locally in the region where they are needed, for the specific application needed, independently from other reactors.

In contrast to fed-batch reactors, the reaction is not controlled by operators adding components over time, but 'in space'. For example, it would be possible to incorporate auxiliary inputs at the ends of specific sections, where components can be added if necessary. The implementation of a robust control system, could allow reactors to be operated automatically in modular, sterile boxes.

5. Conclusion

Mechanistic modelling appears to be a viable alternative for the implementation of quality by design to applications where data-driven modelling is inhibited by costly experiments. Utmost importance is placed on developing a solid process and product understanding instead of only optimising process conditions and illustrating CQA and CPP correlations. By raising the question of why high NTP^{4-} concentrations gives low RNA yield, a potential cause-effect relationship is introduced for further experimental investigation. However, care must be taken when extrapolating the model to unexplored regions that the model was not fitted to, since mechanistic pathways can change, as shown by the failure of the mechanistic model to accurately predict that the RNA yield is essentially zero at high Mg^{2+} and NTP^{4-} concentrations. When implementing mechanistic design spaces, it is also crucial to have a robust way of quantifying uncertainty, which is a drawback of mechanistic modelling compared to multivariate statistical modelling. Finally, even though it is shown that mechanistic modelling can successfully be used in creating a design space for RNA vaccines, this work brings up more questions than it answers, mainly in the realm of physical understanding of biological systems.

From a computational perspective, different numerical methods can be implemented to improve the robustness of the current dynamical model. For instance, a specialised least-squares minimisation solver needs to be implemented to find the global optimal parameters of the system of differential-algebraic equations. Additionally, instead of solving the system of differential-algebraic equations explicitly, one could implement a special solver for high-index differential-algebraic problems. Furthermore, additional data and more robust data validation is needed to build on the current model, implement additional components as CPPs, and ideally link each CQA to all CPPs. This is currently hindered by a major lack of mechanistic understanding of many biological phenomena such as accuracy of sequence identity. This opens up opportunities for the design of hybrid mechanistic, stochastic and machine learning models to extract the best of each framework in terms of cause-effect relationships, risk management and predictive power.

Most importantly, the current model lays down the groundwork for process development, optimisation and control. For instance, one could look for the optimal normal operating range within the design space. Finally, one could link measurable variables to manipulatable process variables. Then, considering current reaction time and RNA yield, one could construct dynamic design spaces to find the CPPs that give the desired final RNA concentration in a first step towards model predictive control.

References:

[1] E. Callaway, "FIRST VACCINE AGAINST DEADLY EBOLA VIRUS WINS APPROVAL," *Nature*, pp. 425–426, Nov-2019.

[2] Z. Kis, R. Shattock, N. Shah, and C. Kontoravdi, "Emerging Technologies for Low-Cost, Rapid Vaccine Manufacture," *Biotechnology Journal*, vol. 14, no. 1. Wiley-VCH Verlag, 01-Jan-2019.

[3] N. Pardi, M. J. Hogan, F. W. Porter, and D. Weissman, "mRNA vaccines-a new era in vaccinology," *Nature Reviews Drug Discovery*, vol. 17, no. 4. Nature Publishing Group, pp. 261–279, 28-Mar-2018.

[4] F. Titta, "Vaccines Drug Product development: how QbD paradigm is the driving principle for product and process understanding."

[5] A. S. Rathore, "Roadmap for implementation of quality by design (QbD) for biotechnology products," *Trends Biotechnol.*, vol. 27, no. 9, pp. 546–553, 2009.

[6] R. E. Baker, J. M. Peña, J. Jayamohan, and A. Jérusalem, "Mechanistic models versus machine learning, a fight worth fighting for the biological community?," *Biol. Lett.*, vol. 14, no. 5, 2018.

[7] FVMR, "Future Vaccine Manufacturing Research (FVMR) Hub," 2019. [Online]. Available: https://www.imperial.ac.uk/future-vaccine-hub/resources/. [Accessed: 11-Dec-2019].

[8] P. M. J. Jedrzejewski, I. J. del Val, K. M. Polizzi, and C. Kontoravdi, "Applying quality by design to glycoprotein therapeutics: experimental and computational efforts of process control," *Pharm. Bioprocess.*, vol. 1, no. 1, pp. 51–69, Apr. 2013.

[9] S. Akama, M. Yamamura, and T. Kigawa, "A multiphysics model of in vitro transcription coupling enzymatic reaction and precipitation formation," *Biophys. J.*, vol. 102, no. 2, pp. 221–230, Jan. 2012.

[10] J. Haas *et al.*, "Implementation of QbD for the development of a vaccine candidate," *Vaccine*, vol. 32, no. 24, pp. 2927–2930, May 2014.

[11] CMC and VWG, "A-VAX: Applying Quality by Design to Vaccines CMC-Vaccines Working Group," 2012.

[12] J. P. Lawrence and J. P. L. London, "MICROFLUIDIC TANGENTIAL FLOW FILTER AND CONTINUOUS-FLOW REACTOR FOR BIOPROCESS DEVELOPMENT," 2015.

[13] Centers For Disease Control and Prevention, "Vaccine Information Statements (VISs) RABIES VIS," 2009. [Online]. Available: https://www.cdc.gov/vaccines/hcp/vis/vis-statements/rabies.html. [Accessed: 12-Dec-2019].

[14] R. H. Garrett and C. M. Grisham, *Biochemistry*. Brooks/Cole, Cengage Learning, 2010.

[15] Y. Li and R. R. Breaker, "Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2γ-hydroxyl group," *J. Am. Chem. Soc.*, vol. 121, no. 23, pp. 5364–5372, Jun. 1999.

[16] M. Oivanen, S. Kuusela, and H. Lönnberg, "Kinetics and Mechanisms for the Cleavage and Isomerization of the Phosphodiester Bonds of RNA by Brønsted Acids and Bases," 1998.

[17] P. Thomen, P. J. Lopez, U. Bockelmann, J. Guillerez, M. Dreyfus, and F. Heslot, "T7 RNA polymerase studied by force measurements varying cofactor concentration," *Biophys. J.*, vol. 95, no. 5, pp. 2423–2433, Sep. 2008.

[18] M. Frugier, C. Florentz, M. W. Hosseini, J. marie Lehn, and R. Giegé, "Synthetic polyamines stimulate in vitro transcription by T7 RNA polymerase," *Nucleic Acids Res.*, vol. 22, no. 14, pp. 2784–2790, 1994.

[19] S. Fjelstrup *et al.*, "The effects of dithiothreitol on DNA," *Sensors (Switzerland)*, vol. 17, no. 6, Jun. 2017.

pH- and Redox-Responsive Hydrogels in Nanoemulsion Systems for Oral Drug Delivery

Abirami Ravindran & Chloe Wong

Abstract—

Oral drug delivery of hydrophobic and super hydrophobic drugs still pose a considerable challenge. To tackle this problem, pH- and redox- responsive hydrogels were combined with nanoemulsion systems to facilitate drug targeting and increase the bioavailability of these drugs. Poly(L-lysine isophthalamide) (PLP) hydrogels were cross-linked with L-lysine methyl ester (CDE) and loaded with oil-in-water (O/W) nanoemulsions. These nanoemulsions contained Nile red or curcumin, and were formulated via spontaneous emulsification, using Eumulgin RH 40 as a surfactant. Successful cross-linking of the hydrogels was verified using FTIR spectroscopy and SEM. The long-term stability of the drug-loaded nanoemulsions was verified using differential light scattering and zeta potential analyses. The pH- and redox- responsive behaviour of the hydrogels allowed the drugs to be protected in a simulated gastric environment, and released in the simulated intestinal fluid, where absorption is maximum.

Index Terms—hydrogel, nanoemulsion, hydrophobic, pH-response, redox-response, Nile red, curcumin

1 Introduction and Background

Oral administration is the preferred way of drug delivery into the body due to convenience, patient acceptance, and ease of use especially in cases of regular dosing (Sant et al. 2005). Hydrogels are porous, hydrophilic networks capable of swelling to absorb great amounts of water or other fluids. They are highly applicable in the biomedical industry for oral drug delivery due their low cost, biocompatibility, and ability to respond to different environmental stimuli, for example temperature, light, and ionic changes (Chatterjee & Hui 2018).

In particular, within the human body the stimuli of greatest interest are pH and redox changes. The hydrogels presented in this paper were first synthesised by Wang et al. (2016). They are formed from a polymer backbone of Poly(L-lysine isophthalamide) (PLP) which is an amphiphilic pseudo-peptide of pKa 4.4 (Eccleston et al. 2000). Below pH 4.4, the hydrogel structure becomes very compact due to coil-to-globule conformational changes and subsequent hydrophobicity variations (Wang, Attah, Li, Chen & Chen 2018). These variations come about from the carboxylic acid groups of PLP protonating. At pH 4.4 and above, the carboxylic acid groups deprotonate and the negative charge causes electrostatic repulsion, creating the swelling behaviour seen in higher pH environments. This leads to a greater drug release.

This is important because there is a notable pH variation along the gastrointestinal tract within the human body. The stomach is very acidic with a pH range of 1.0-2.5 whereas the colon (or large intestine) has more alkaline conditions with a pH range of 6.6-7.5 (Evans et al. 1988). The pH-responsiveness of these hydrogels allows for localised drug delivery to the colon. The colon provides an ideal enivronment for drug release as it has a large surface area, long retention time, less variation and intensity of digestive enzymes, leads to a better bioavailability, and is highly responsive to absorption enhancers (Philip & Philip 2010, Kumar et al. 2010).

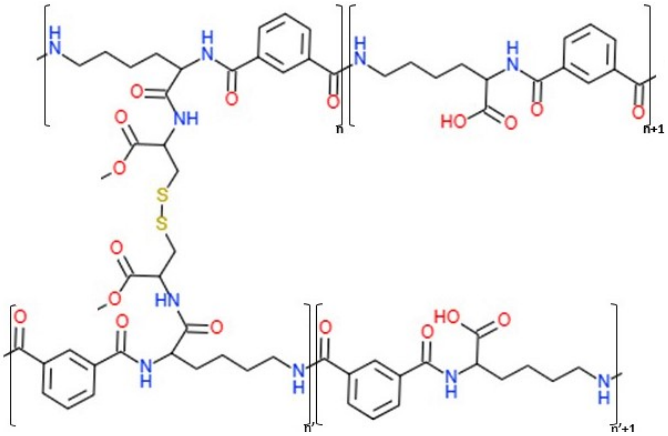


Fig. 1: Hydrogel structure: PLP crosslinked with CDE

The PLP hydrogels were crosslinked with L-cystine dimethyl ester dihydrochloride (CDE), an amino acid derivative, via an N-(3-dimethylaminopropyl)-N'ethyl carbodiimide hydrochloride(EDC)/N-hydroxysuccinimide (NHS) coupling reaction. The final structure is shown in Figure 1. Crosslinkers are hugely important for the structure of hydrogels. It creates insolubility of the polymer due to strong covalent bonds, but still allows the hydrogel to absorb water. They also introduce hindrance to chain mobility therefore decreasing the viscosity and improving the stability of the hydrogel (Maitra & Shukla 2014).

CDE allows for the incorporation of redox-responsiveness through breakage of its disulphide bond. The redox potential within the colon has been measured to be -415 $\pm$ 72 mV which is sufficient to cleave the disulphide linkage, therefore leading to a more complete release of payload than with pH-response alone (Stirrup

et al. 1990). 1,4-dithiothreitol (DTT) was used as a redox agent to reduce the disulphide bonds and replicate the redox conditions in the colon.

The number of disulphide bonds can be varied to give optimal degradation properties of the hydrogel (Wilding et al. 1994). In addition, the amount of crosslinker itself can be changed to study its effects on hydrogel characteristics and pH-response.

It is estimated that at least 40 % of marketed drugs and up to 60 % of drugs currently in research and development stages are poorly water-soluble, therefore the need for hydrophobic drug carriers is ever more pressing (Fahr et al. 2007). Hydrogels are intrinsically hydrophilic but there is significant research into adapting them for hydrophobic drug use.

Other systems have also been investigated for hydrophobic drug delivery such as micro- and meso-porous silica carriers (Bremmell & Prestidge 2019) and polymeric micelles (Xu et al. 2013), however the main focus remains on hydrogels. Examples of some of the research that has been carried out include Watkins & Chen (2015) who investigated pH-responsive, amphiphilic PLP hydrogels crosslinked with L-lysine methyl ester displaying the ability to deliver payloads greater than 70 kDa, and Martin et al. (2003) who prepared amphiphilic GCP hydrogels via physical crosslinking.

Despite there being much research on the delivery of hydrophobic drugs, the challenge still remains with the delivery of drugs such as Nile Red, a model drug that is super hydrophobic. Wang, Attah, Li, Chen & Chen (2018) formed pH-responsive amphiphilic hydrogels using PLP crosslinked with PEG, however Nile red still had a low loading capacity. Wang, Ha, Huang, Chin, Sim & Chen (2018) developed novel pH-responsive and membrane-active nanogels to carry Nile red, taking advantage of both the size of nanotechnology and the flexibility of hydrogels. The drug loading and release profiles obtained were promising, highlighting the potential of nanosized systems.

Nanoemulsions are another novel system that has yet to be investigated thoroughly for hydrophobic drug release. They are colloidal particulate systems formed from two immiscible liquids, one closely dispersed in the other, that are stabilised by a surfactant (Jaiswal et al. 2014). This paper investigates an oil-in-water (O/W) nanoemulsion loaded into PLP-CDE hydrogels as a hydrophobic drug delivery route.

There are numerous advantages to the use of nanoemulsions including non-toxicity. An O/W system not only addresses the challenge of transporting super hydrophobic or lipophilic drugs, it also envelops and protects the carrier drugs from enzymatic degradation and hydrolysis (Lovelyn & Attama 2011), which may make it a suitable option for the transport of sensitive drugs. Due to the nanosized nature of the system, a very large interfacial area provides better absorption and allows for drug targeting, which leads to an increased bioavailability of carrier drugs (Kim et al. 2001, Wagner et al. 1966). According to Farokhzad & Langer (2009) and Lavasanifar et al. (2002), nanoemulsions also allow for transcytosis of drugs across the tight intestinal barrier.

In addition, there are low-energy methods to formulate these nanoemulsions. As demonstrated by Shakeel et al. (2008), phase inversion temperature is a simple formulation technique that involves employing a higher temperature to a microemulsion. Spontaneous emulsification requires energy only for continuous magnetic stirring, and was applied in this paper to formulate O/W nanoemulsions. There are also high-energy methods that typically require specialised equipment such as high-pressure homogenisers, microfluidsers and ultrasonicators (Jaiswal et al. 2014).

The role of the surfactant in these systems is to provide thermodynamic stability by enhancing zeta potential and viscosity. The selected surfactant should be able to reduce the surface tension while being fluidic enough to produce a coherent film that prevents coalescence. Ideally, a surfactant should be effective in small concentrations (Nikam et al. 2018, Grumezescu 2019). There is much about the role of the surfactant that still needs to be understood.

There has been some interest in nanoemulsion systems containing curcumin due to curcumin's antioxidant, anti-inflammatory, antimicrobial and anticancer properties (Sari et al. 2015). Despite its many benefits, the current pharmaceutical applications of curcumin are limited by its hydrophobicity, rapid metabolism, poor absorption and low bioavailability in the gastrointestinal track (Sood et al. 2014). Studies concerning parenteral curcumin nanoemulsions have looked at intranasal delivery targeted at the brain, to overcome blood–brain and blood–cerebrospinal fluid barriers (Sood et al. 2014); and transdermal applications where the nano size of the system allowed for improved permeation through skin pores (Rachmawati et al. 2015). Wang et al. (2008) used high-pressure homogenisation to formulate curcumin O/W nanoemulsions, while Cui et al. (2009) investigated self-microemulsifying drug delivery systems (SMEDDS). This paper will aim to tackle the oral delivery of curcumin, formulating nanoemulsions using a low-energy method while avoiding SMEDDS as they typically require a high amount of surfactant, in the case of Cui et al. (2009), the optimal SMEDDS contained 57.5 % surfactant.

High-pressure homogenisation was used to formulate paclitaxel-containing nanoemulsions by Kim & Park (2017) and paclitaxel-loaded lipid nanoemulsions were prepared by Chen et al. (2017) for targeted tumour delivery. However, there is still much research to be done regarding super hydrophobic drugs, therefore in addition to curcumin, Nile red drug loading was also examined.

2 Experimental Section

2.1 Materials

Obtained from Sigma Aldrich (UK): Eumulgin RH 40, Oleic acid, L-cystine dimethyl ester dihy-

drochloride (CDE), N-hydroxysuccinimide (NHS), N-(3-dimethylaminopropyl)-N'ethyl carbodiimide hydrochloride (EDC), 1,4-dithiotheritol (DTT), Nile Red, Curcumin.
Obtained from VWR International Ltd. (UK): Absolute ethanol, Disodium phosphate (Na_2HPO_4), Monosodium Phosphate (NaH_2PO_4), Sodium hydroxide (NaOH), Hydrochloric Acid (HCl).

Obtained from Thermo Fisher Scientific (UK): Citric acid, Sodium chloride (NaCl), Monopotassium phosphate (KH_2PO_4).

Poly(L-lysine isophthalamide) (PLP) was obtained in its acidic form from Michal Kopytynski and in its salt form from Apanpreet Bhamra. All materials were used as received, apart from the acidic form of PLP, which was neutralised and freeze dried before use.

Methods

2.2 Synthesis of pH and redox-responsive hydrogels

Hydrogels were synthesised as previously reported by Wang (2019). To make two hydrogels, 100 mg PLP was dissolved in 1 mL of deionized (DI) water. Under constant mixing with a magnetic stirrer, varying amounts of CDE (Table 1) was slowly added to the polymer solution. After the CDE had dissolved, 140 mg EDC, 20 mg NHS, and an extra 1 mL DI water was added. This was stirred until everything had dissolved and then transferred to a 48-well plate, where 1 mL of solution was pipetted into each well. The final solution was left for up to 48 hrs until gelation occurred. After gelation, the hydrogel was transferred into a beaker of DI water to remove impurities. The water was changed daily for at least three days.

TABLE 1: CDE amount used in the synthesis of different hydrogels

Hydrogel Number	CDE amount (mg)
1	23
2	26
3	30
4	33

2.3 Buffer Solution Preparation

2.3.1 Simulated Intestinal Fluid (SIF)

1.7 g of KH_2PO_4 and 0.225 g of NaOH were added into 250 mL of deioninsed (DI) water. The pH was then read and adjusted to 6.8 using 0.1 mol/L NaOH buffer or 0.1 mol/L HCl buffer. SIF was used to imitate how the drug may be absorbed in the intestine. 0.125 mol/L DTT was added to SIF buffer to test the redox response of the hydrogels.

2.3.2 Simulated Gastric Fluid (SGF)

0.5 g of NaCl and 1.75 mL of HCl were added to 250 mL DI water. The pH was then adjusted to 1.2 to mimic the conditions in the stomach.

2.3.3 Citrate buffer

100 mM of citrate buffer was used to discharge any water left in the hydrogel from purification. It was prepared by adding 0.588 g of citrate powder into 250 mL DI water. The pH of the buffer was adjusted to 3.0.

2.4 Nanoemulsion Synthesis

Spontaneous emulsification was used to formulate the O/W nanoemulsion.

2.4.1 Oil Phase

The oil phase was made by adding 18.5 g of RH40 and 5 g of oleic acid into 76.5 g of ethanol. The solution was stirred until homogenous via magnetic stirring. The hydrophobic drugs were measured out and loaded into the oil phase under continuous magnetic stirring. Drug concentrations to be loaded were selected as 0.3 mg/mL and 3 mg/mL of Nile red and curcumin respectively in ethanol.

2.4.2 Water Phase

The water phase was prepared by adding 535.9 mg of Na_2HPO_4 and 147 mg of NaH_2PO_4 into 250 mL DI water. The pH of the solution was adjusted to 7.4 using 0.1 NaOH buffer or 0.1 HCl buffer.

The oil phase was then injected into the water phase. Magnetic stirring was applied continuously during this process and continued after for a period of 12 hours.

2.5 Drug Loading

3 mL of citrate buffer was added into a tube containing one hydrogel. The hydrogel was left to soak in buffer for 48 hours to shrink the hydrogel and discharge water in order to increase encapsulation. The citrate buffer was removed and the hydrogel rinsed with DI water before being soaked in 10 mL of drug-loaded nanoemulsion for 48 hours. 5.4 μL of 14 % (v/v) HCl buffer was then added and the hydrogel left for a further 24 hours to closely pack the drug in the hydrogel.

2.6 Drug Release

The drug loaded hydrogel was dropped into 10 mL of SIF, SGF or SIF + DTT buffer and placed in a water bath at 37.5 °C with gentle agitation applied. 150 μL of sample was taken out every hour for 6 hours then once at 24 and 48 hours each. For every sample taken, 150 μL of fresh buffer was put back in the tube.

Characterisation Techniques

2.7 Fourier Transform Infrared (FTIR) Spectroscopy

Fourier Transform Infrared spectroscopy allows for the characterisation of hydrogels as each bond absorbs a different wavelength of IR radiation, thus producing a spectra. Before FTIR, pure hydrogel samples were placed in a -80 °C freezer for over 3 hours then freeze-dried for 24 hours.

2.8 Scanning Electron Microscopy (SEM)

Scanning electron microscope was utilised to obtain pictures of surface morphology variations in hydrogels of different CDE amounts and in different pH's. Samples were prepared by soaking for 48 hours in buffers to equilibrate. Each hydrogel was placed in SGF buffer (pH 1.2) and SIF

buffer (pH 6.8) so the pH-response could be studied. After this, the samples were lyophilised for over 24 hours. This is because inside the SEM is a high vacuum environment so only dehydrated samples can be used. Before entering the SEM, tiny pieces of lyophilized hydrogel were attached to conductive tape and coated in a thin layer of gold by an EMITECH K550 coater for 2 minutes with a current of 20 mA. Images were then obtained using a JEOL6010LA microscope with an accelerating voltage of 20 kV.

2.9 Swelling Study

Swelling studies enabled the examination of the pH and redox-response of the hydrogels and how this changed with respect to the amount of CDE used. As reported in Wang (2018) and Wang (2019), PLP-CDE hydrogels expand in high-pH environments allowing more drug to be released, and contract in low-pH environments, keeping the drug tightly packed inside the hydrogel.

All hydrogels were soaked in 0.1 M phosphate buffer at pH 7.4 for 48 hours at room temperature. They were then equilibrated in different buffers for another 48 hours. The buffers utilised were 0.1 M citrate buffer (pH 3), stimulated intestinal fluid buffer (pH 6.8), stimulated intestinal fluid buffer + DTT (pH 6.8), and stimulated gastric fluid buffer (pH 1.2). After this stage, the hydrogels were weighed to obtained their wet weight, m_w.

To obtain the dry weight, m_d, the hydrogels were soaked in DI water for 48 hours and lyophilised before being weighed.

Error bars were calculated using the standard error formula:

$$SE = \frac{\sigma}{\sqrt{n}} \tag{1}$$

where σ is the standard deviation of the sample set and n is the number of samples taken.

2.10 Differential Scanning Calorimetry (DSC)

Differential scanning calorimetry was used to detect thermal transitions in hydrated hydrogels.

To prepare the samples, 3-5 mg of pure hydrogel samples were weighed in Tzero Aluminium Hermetic pans and sealed. They were then placed in a TA Q200 DSC and tested by heating to 80 °C and then cooling to -60 °C for a total for three cycles. The ramp rate was 10 °C/min.

2.11 Dynamic Light Scattering (DLS)

Differential light scattering was used to measure particle size and particle size distribution in the nanoemulsion. Several nanoemulsion samples were prepared for DLS including pure nanoemulsion as well as nanoemulsion loaded with 0.3 mg/mL Nile red and 3 mg/mL curcumin. The long-term stability of the drug-loaded nanoemulsions were tested by taking samples once every 4 days, over a period of 20 days. Characterisation of the nanoemulsion released from hydrogels was carried out by taking samples released into SIF, SGF and SIF + DTT buffers after 48 hours. 2 mL of each sample was collected and filtered with a 200 nm syringe filter, then dropped into a semi-micro UV cuvette. The samples were measured at 25 °C, at an angle of 90 ° and fitted by Malvern software.

2.12 Zeta Potential

The Brookhaven ZetaPALS Potential Analyzer was used to determine the stability of the nanoemulsion by calculating the zeta potential. If particles have a large negative or positive zeta potential, usually taken as less than -30 mV or larger than 30 mV, then the system is normally considered stable as particles will repel each other (Malvern Instruments Ltd. n.d.). However, if the zeta potential is smaller, there is not enough repulsive force and the particles may agglomerate.

Nanoemulsion samples were prepared by filtering 2 mL samples of pure nanoemulsion as well as nanoemulsion loaded with 0.3 mg/mL Nile red and 3 mg/mL Curcumin with a 200 nm syringe filter. Each sample went through the ZetaPALS Zeta Potential Analyzer for 6 runs and 20 cycles using the Smoluchowski model.

2.13 UV Absorption Spectroscopy

UV Absorption Spectroscopy was used to measure the amount of drug loaded. The concentrations of drug released into SIF, SGF and SIF + DTT buffers were also tested every hour for six hours, then once at 24 and 48 hours. The absorbance of the loading solutions before and after loading, as well as the absorbance of the release buffers were tested at 405 nm and 490 nm for Nile red and curcumin respectively, via an ultraviolet-visible light (UV-Vis) spectrophotometer (Genesys 10S). 150 μm of each sample was dropped into a well on a 96-well plate. Concentrations of the drug in the samples was then determined via correlation curves of Nile red and curcumin in the same conditions, at the aforementioned wavelengths. The encapsulation efficiency, loading efficiency and percentage of drug released were calculated using the following relations:

$$\text{encapsulation efficiency } (\%) = \frac{m_l}{m_t} \times 100 \tag{2}$$

$$\text{drug released } (\%) = \frac{m_r}{m_l} \times 100 \tag{3}$$

where m_l is the mass of drug loaded into the hydrogel, m_t the total mass of drug in the loading solution, and m_r the amount of drug released at a specified time.

3 Results and Discussion

3.1 Hydrogel Synthesis and Characterisation

For all hydrogels, PLP and CDE cross-linker were added into a beaker, followed by EDC and NHS to facilitate the coupling reaction. The cross-linking reaction occurred between the carboxylic acid groups of PLP and the amine groups of EDC as shown in Figure 1. After mixing, the clear liquid was pipetted into a 48-well plate and left to gelate over night, forming a solid, white, opaque hydrogel. The

rate of gelation was observed to be faster the greater the amount of CDE, with Gel 4 turning a milky colour even during the mixing stage.

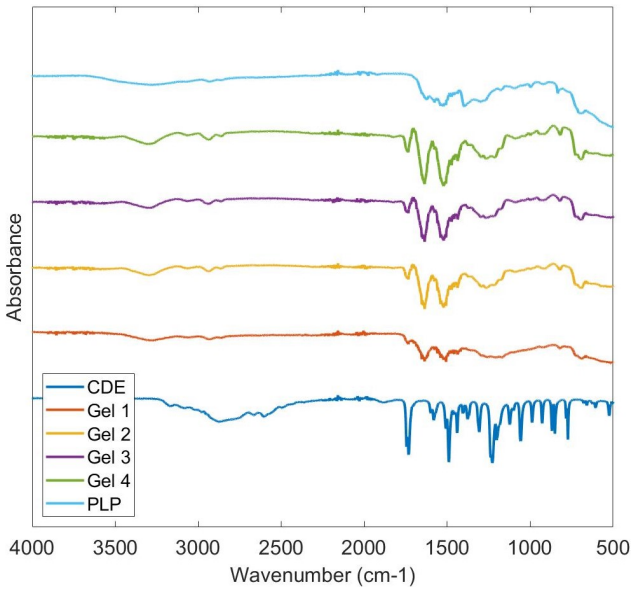


Fig. 2: FTIR Spectra for different hydrogels, PLP, and CDE

FTIR spectroscopy was utilised to confirm the successful fabrication of PLP-CDE hydrogels. All CDE amounts lead to very similar looking spectras. Characteristic peaks attributed to PLP can be seen at 1636 cm^{-1}, which is the amide I band for C=O stretching, directly related to the polymer backbone conformation (Jabs n.d.), and 1522 cm^{-1} from the amide II band for N-H stretching. A CDE characteristic peak was seen at 1734 cm^{-1} for the C=O stretching of the ester group. The increased intensity of the broad peak at around 1229 cm^{-1} in the hydrogel structure can probably be attributed to the strong, sharp peak in the CDE spectra at the same wavelength. This could be the C-N stretching bond. The disulfide S-S stretch and C-S stretch of CDE should be seen at wavelengths between 700-550 cm^{-1} and 710-570 cm^{-1} respectively, however they have weak IR transmittance and have no visible effects on the hydrogel spectras (Hampton et al. 2010).

3.2 Hydrogel Morphology and Porosity

The morphology and porosity of hydrogels are important as the number of pores and pore size is ultimately what determines the amount of drug that can be loaded and unloaded.

SEM was used to study these qualities. As seen in Figure 3, there was a large variation in the morphology between the different gels when placed in SIF buffer. Gel 1 looked flat with ridges across the surface, Gels 2 and 4 had a much more porous structure and a more uneven surface, and Gel 3 also had a porous structure but with less definition of pore outlines. There wasn't a clear trend in pore size or structure between the gels, which is due to differences in morphology even within the same hydrogel depending on factors such as whether the sample was taken from the outside or inside of the hydrogel.

Gel 1 had far fewer and much less defined pores than the rest of the hydrogels. This can be explained as Gel 1 had a lower crosslinking density therefore forming a less robust hydrogel, so it is possible that during the lyophilisation process the hydrogel network collapsed.

It was difficult to measure the exact pore size of the gels from the SEM pictures obtained, however the rough pore size of Gel 2 is estimated to be 60 μm based on the upper left indentation. As the amount of crosslinker increases, the average pore size is expected to decrease due to the hydrogel network become more densely packed.

When placed in SGF buffer, the SEM pictures taken showed different morphologies, with little to no pores visible. Some displayed pitting but with very poor definition between the pits and whilst others had a sand-like quality to them. Different morphology is expected with low pH conditions as the hydrogels should have contracted, leaving little to no pores.

3.3 Hydrogel pH- and Redox-Responsiveness

A swelling study, displayed in Figure 4 was carried out to test the pH- and redox-responsiveness of the hydrogels. The swelling ratio, q is defined as:

$$q = \frac{m_w}{m_d} \tag{4}$$

It is an indicator of how much a hydrogel's network can expand in a certain environment, and thus how much liquid and drug can be loaded inside.

As expected, there was good pH- responsiveness for all gels tested. Both SGF and citrate buffer gave low q values in the approximate range of 5-10 whilst in pure SIF the q values increased to about 20. The similarity in results for the SGF and citrate buffer can be attributed to the pH's all being less than the pKa value for PLP of 4.4.

The change in swelling ratio a high and low pH environment was also examined using the formula:

$$\Delta q = q_{SIF} - q_{SGF} \tag{5}$$

TABLE 2: Change in swelling ratio for different hydrogels from SGF to SIF buffer

Hydrogel Number	Δq
1	14.8 ± 0.5
2	10.1 ± 1.0
3	7.6 ± 0.2
4	10.6 ± 0.8

Table 2 shows the values of Δq obtained. It can be see that between Gels 2, 3, & 4 there wasn't a huge difference between the values, however Gel 1 had a much higher Δq in comparison with a value of 14.8. As stated previously, Gel 1 had lower amounts of CDE crosslinker so there was a less dense hydrogel network, allowing for an increase in swelling

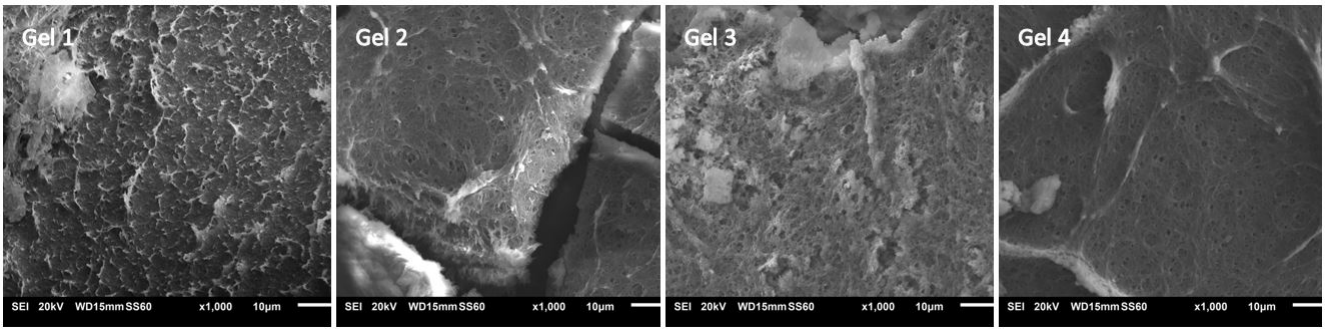


Fig. 3: SEM photographs of different hydrogels in SIF buffer (pH 6.8)

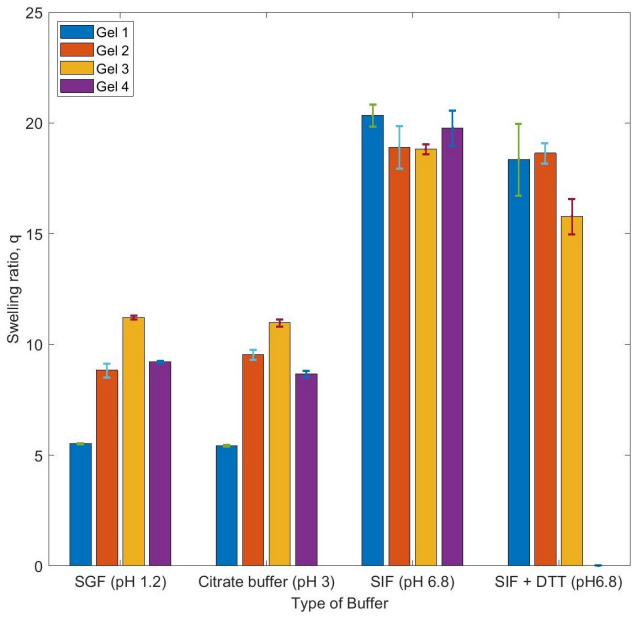


Fig. 4: Swelling Study for hydrogels in different pH and redox environments

compared to the other hydrogels. Furthermore, less CDE leads to more free carboxylic acid groups on the polymer backbone and therefore an increase in pH activity.

The redox-responsiveness was tested with the addition of DTT into the SIF buffer. Gel 4, being the gel with the most CDE and disulfide bonds to break, was degraded much more than the other gels, to the extent that q could not be calculated because it could not be removed from the tube. The results obtained for the other gels were counter-intuitive, with q either staying the same or decreasing compared to the SIF buffer values, even though a visible increase in swelling was seen (see supporting information). This may be due to the degraded hydrogel not being able to absorb as much water, and therefore having a lower wet weight. For qualitative results, the swelling study was able to pick up on differences between the buffers by visual examination. However, for quantitative results, a swelling study is not the best experiment to perform to test redox behaviour due to inaccuracies in correlating swelling to redox-responsiveness.

Previous research by Wang et al. (2016) into the mechanical properties of hydrogels also gave results that match the degradation and redox-response observed. Dynamical rheological measurements were carried out to determine the storage modulus, G', and the loss modulus, G" at different pHs. If G' $>$ G" the material is mainly elastic (Franck n.d.). Wang measured G' for the hydrogels soaked in DTT to be only 1% of G' at pH 3.0 suggesting the hydrogel in DTT had completely dissociated. Furthermore, G" became greater than G' at a higher frequency with the addition of DTT than for a hydrogel at the same pH without DTT, so the addition gave a more liquid-like consistency (Wang et al. 2016).

The standard error was larger for the more swelled hydrogels, likely due to slight variations in pore size within the same gel category and thus different uptakes of water.

3.4 Hydrogel Thermal Analysis

Thermal analysis of the hydrated hydrogels was undertaken using DSC as shown in Figure 5. The DSC peaks represent the transitions of non-freezing bound water, freezing bound water and freezing free water (Finosh 2013). At 0 °C there is the endothermic peak of melting freezing free water and melting freezing bound water. The peak occurs at the same place for all the hydrogels. Between, -10 to -20 °C a large exothermic loop representing the crystallisation of freezing water. The loop is a result of the crystallisation process. Before crystallisation begins, the sample undergoes supercooling and then as exothermic process of crystallisation starts the temperature rises again. Once crystallisation has completed, the temperature drops again therefore leading to the loop seen in Figure 5 (Höhne 2019).

The crystallisation loop shifts to higher temperatures as the CDE content increases as shown in Figure 6. This suggests that the degree of crosslinking has an effect on water state and behaviour. One hypothesis is that a greater amount of CDE leads to a more dense network and therefore a lower water content. As there are less water molecules, not as much energy is needed to overcome the kinetic energy of the molecules and fix them into their crystalline conformation, thus leading to a higher crystallisation temperature.

3.5 Nanoemulsion Particle Size

Chuacharoen et al. (2019) found that adjusting surfactant concentration resulted in changing nanoemulsion

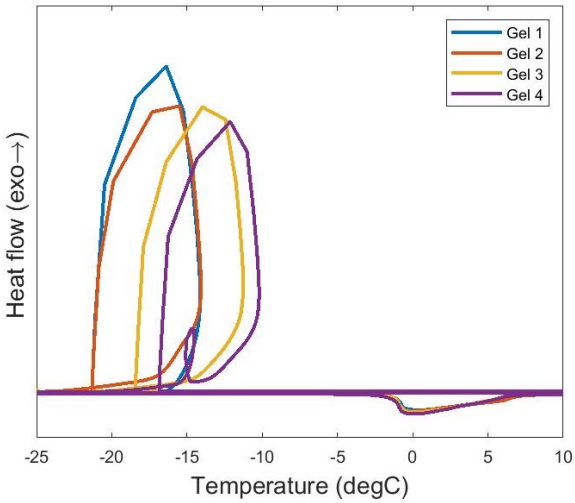


Fig. 5: DSC curves for different hydrogels

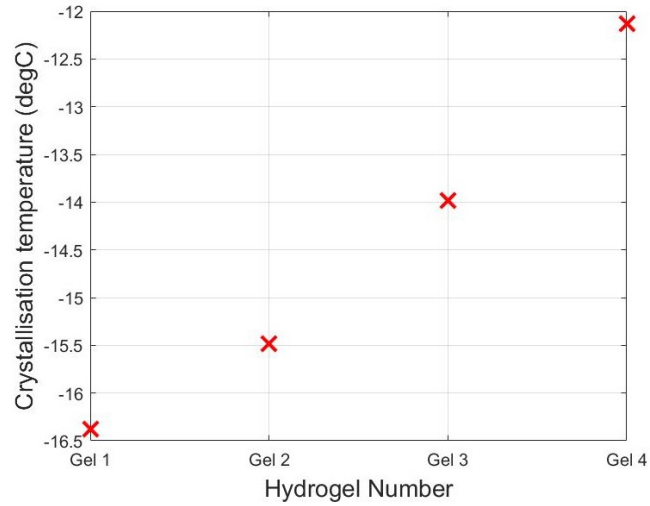


Fig. 6: Temperature of water crystallisation for different hydrogels

particle sizes. This provides a simple method to modify nanoemulsion particle sizes to desired values, allowing successful loading into hydrogels. The pure O/W nanoemulsion was found to have a droplet size of 16.0 ± 1.7 nm. Nile red and curcumin O/W nanoemulsions had average particle sizes of 17.4 ± 2.8 nm and 18.3 ± 2.9 nm respectively. The increased standard deviation upon loading drug into the nanoemulsions could suggest that additives affect the surface tension of the system.

Nanoemulsion particle sizes are generally expected to fall around 10 – 500 nm (Sharma et al. 2013). The sizes reported in the previous paragraph fall within the expected range which indicates that the nanoemulsion was successfully loaded into hydrogels. They also fall on the lower end of expected values which indicates larger overall surface area that will in turn give rise to more efficient absorption in the intestine during drug release.

Figure 7 indicates that the nanoemulsion particle sizes remained similar before and after drug release into SIF, SGF and SIF + DTT buffers. This implies that the drug loading process did not damage or break the nanoemulsion particles which is favourable. Additionally, despite contact with a range of pH conditions (from pH 1.2 to pH 6.8) and redox agents, the nanoemulsion was still able to protect the carrier drug. The indication that the carrier drugs remained enveloped within the nanoemulsion throughout the drug release process provides us with a feasible route for the transfer of sensitive or super hydrophobic drugs.

3.6 Nanoemulsion Stability

3.6.1 DLS

The stability of the nanoemulsions was verified via zeta potential and particle size measurements over 20 days. As per Figure 8, the particle sizes of both Nile red and curcumin loaded nanoemulsions remain relatively constant over this period. There appears to be a very slight increase in the trend, which might be attributed to the fact that surfactants may start to slowly precipitate back out over time. However, the study should be carried out over a longer period of time to determine if this is indeed the case, or if it is simply a fluctuation in readings. For the most part, the selected surfactant did not appear to decompose. There was also no significant change in the polydispersity index (PDI) for both Nile red and curcumin nanoemulsions, with values remaining below 0.2 and 0.4 respectively. As such, concentrations of 0.3 mg/mL and 3 mg/mL of Nile red and curcumin respectively in ethanol were determined to be suitable for storage and transport.

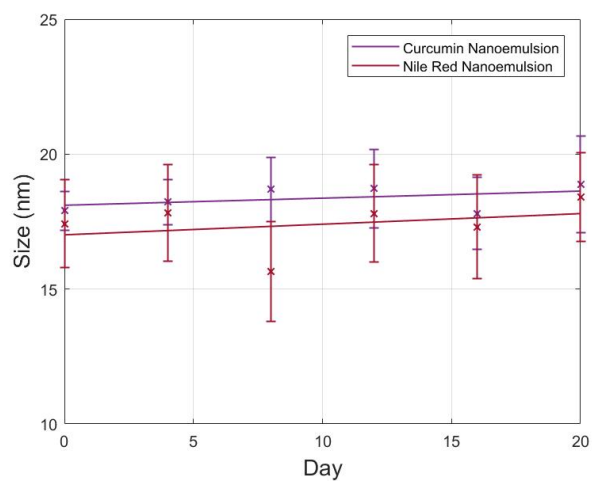


Fig. 8: Particle size analysis for drug loaded nanoemulsions over 20 days

3.6.2 Zeta Potential

The role of emulsifiers further extends to creating a surface potential within the nanoemulsion system. The resultant repulsive forces between oil droplets will prevent agglomeration and ensure the stability of the system, as reported by Kelmann et al. (2007). The electrophoretic mobility measurements recorded were used to calculate zeta potential via the Smoluchowski relation. As all zeta potential values were determined to be more negative than -30 mV, as per Table 3, all the tested systems can be classified as stable for storage.

Due to time constraints, measurements were taken 20 days after the formulation of the nanoemulsions. Measurements should be repeated after a few months to determine

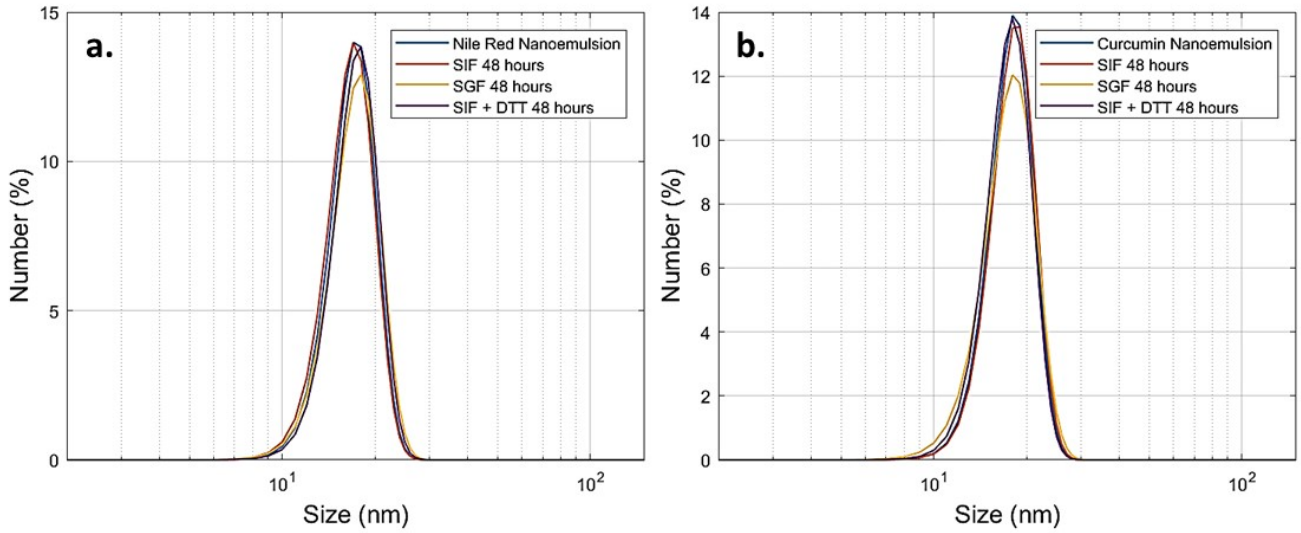


Fig. 7: Nanoemulsion particle sizes before and after 48 hours in release buffers: (a.) analysis on Nile red nanoemulsion, and (b.) analysis on curcumin nanoemulsion

the long-term stability of the system, as this will be dependent on whether the emulsifier precipitates back out over time.

TABLE 3: Zeta Potential and Mobility Values

Nanoemulsion	Mean Zeta Potential (mV)	Mobility (u/s) / (V/cm)
Pure O/W	-38.52 ± 0.12	-3.01 ± 0.01
Nile red	-38.19 ± 0.06	-2.98 ± 0.01
Curcumin	-37.94 ± 0.15	-2.96 ± 0.01

3.7 Drug Loading & Encapsulation

Gel 2 was selected for drug loading as it was deemed to be a good trade-off between pH- and redox- responsive behaviour. Figures 9(a.) and 9(b.) indicate the encapsulation efficiencies, as calculated by Equation 2, of hydrogels in different concentrations of drug-loaded nanoemulsions. Although lower concentrations have higher encapsulation efficiencies, the amount of drug loaded is ultimately lower, and so a compromise must be made.

Figures 9(c.) and 9(d.) indicate that an optimum loading amount is reached at concentrations of 0.3 mg/mL Nile red and 3 mg/mL curcumin in ethanol, after which the amount loaded plateaus off. This is an expected result. When formulating the nanoemulsions, concentrations of up to 1 mg/mL and 10 mg/mL of Nile red and curcumin in ethanol were made as these are the respective solubilities of the carrier drugs in ethanol. The drugs appeared to dissolve in the oil phase at maximum solubility concentrations. However, when injecting the oil phase into water phase, some drug appeared to precipitate back out for higher concentrations. At concentrations of 0.4 mg/mL Nile red and 4 mg/mL curcumin, a substantial amount of carrier drug was observed to deposit back out of the nanoemulsion after 12–24 and 4–6 hours respectively, resulting in a waste of raw materials. As such, 0.3 mg/mL Nile red and 3 mg/mL curcumin were selected as the most economically feasible and sustainable options.

3.8 Drug Release

The pH- and redox-responsive properties of the hydrogels were verified by the release profiles obtained in Figures 9(e.) and 9(f.). Errors were calculated using Equation 1. For both model drugs, there was very little leakage (around 5 % to 6 %) of drug into SGF buffers. This is expected, as per SEM and swelling study analysis, the hydrogel appeared to contract and have very few pores when placed in the SGF buffer. The pH-responsive hydrogels expanded in the colon mimicking environment, releasing 40 % of encapsulated Nile red and 32 % of encapsulated curcumin after 48 hours. The effect of the redox-response was also clear as a further increase of 46 % of Nile red released was observed upon adding DTT. Drug release of curcumin in SIF + DTT could not be completed due to time constraints. Despite the fact that nanoemulsion particle sizes were much smaller than the pore sizes of the hydrogels in SIF buffer, drug release percentages were still less than half the amount encapsulated. This could be due to the complex, cross-linked structure of hydrogels that may hinder the release of drug encapsulated deep within it. Addition of DTT triggers the destruction of the hydrogels, which explains the much more complete release. Combining the nanoemulsion and hydrogel systems tackles previously identified problems regarding the loading of hydrophobic drugs into hydrogels, thereby increasing the bioavailability of these drugs.

4 Conclusions and Outlook

In summary, the amount of crosslinker within the hydrogel structure has a profound effect on the pH- and redox-responsiveness. There is an optimum amount of crosslinker that needs to be used as trade-off between pH- and redox-responses. From the amounts tested in this paper, 26 mg of CDE (Gel 2) was selected as the optimum amount to load

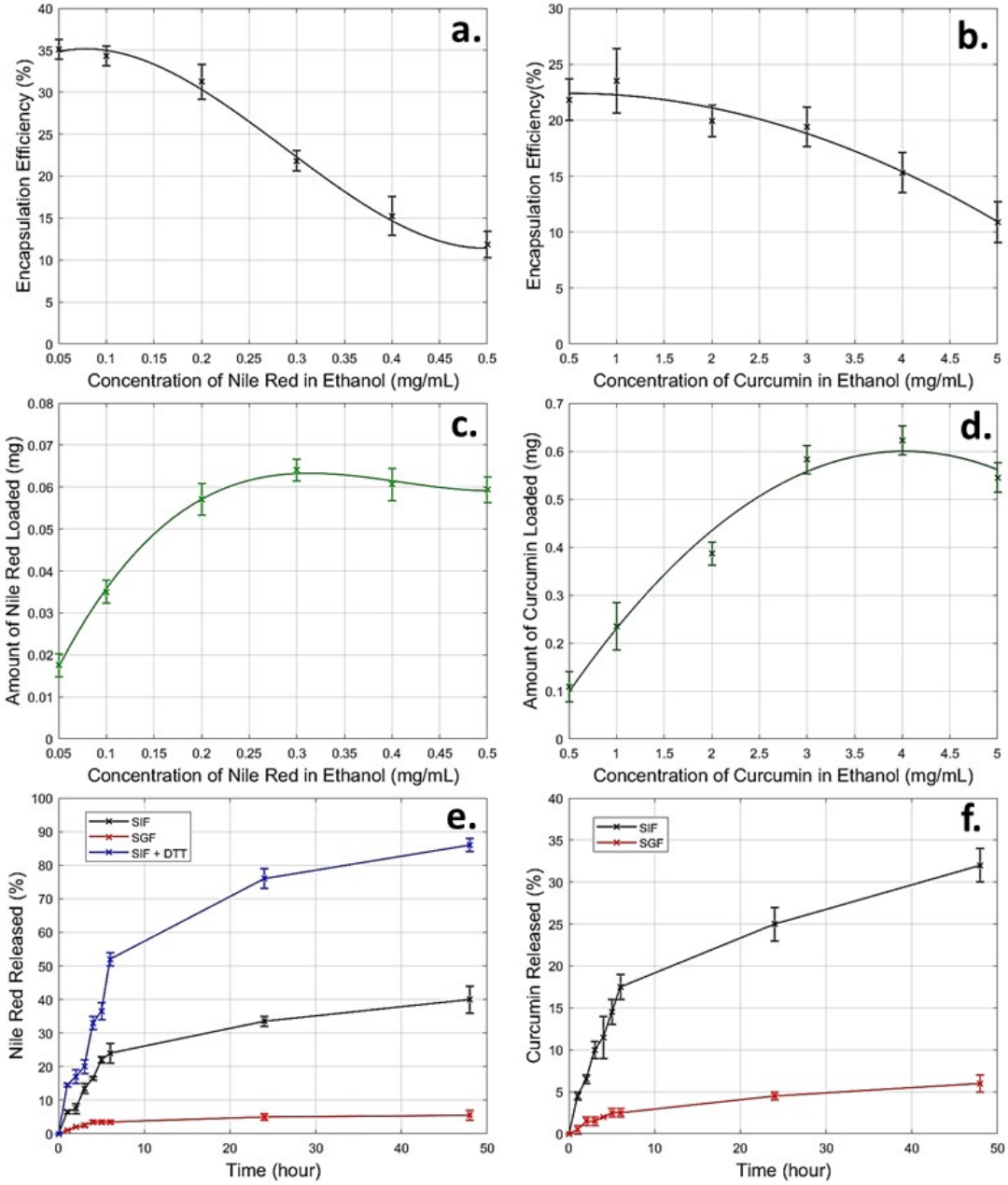


Fig. 9: Drug loading and release profiles for Gel 2: (a.) encapsulation efficiency of hydrogels soaked in different concentrations of drug-loaded Nile red nanoemulsion, (b.) encapsulation efficiency of hydrogels soaked in different concentrations of drug-loaded curcumin nanoemulsion, (c.) amount of drug loaded into hydrogels soaked in different concentrations of drug-loaded Nile red nanoemulsion, (d.) amount of drug loaded into hydrogels soaked in different concentrations of drug-loaded curcumin nanoemulsion, (e.) drug release of Nile red in different releasing buffers, and (f.) drug release of curcumin in different releasing buffers

the nanoemulsion. However, future work could investigate the same CDE range outlined in this paper but with smaller increments to determine the exact optimum. Furthermore, a sufficient amount of crosslinker should be chosen so that enough pores are formed, but excessive quantities and pores formed may be too small. Drug release within this range could also be useful as different amounts for crosslinking would likely change the release profile as pore sizes vary.

The nanoemulsion-loaded systems show promising results for use with super hydrophobic drugs. Both Nile red and curcumin drug release profiles show a big contrast between SIF buffer and SGF buffer giving the desired pH-response, whilst Nile red had almost 90% drug release after 48 hours with the addition of DTT, confirming degradation of the hydrogel. The optimal amounts of drug to load were found to be 0.3 mg/mL for Nile red and 3 mg/mL for curcumin. A path for further research would be to investigate the effects of additives as this could lead to the possibility of even more drug being loaded inside the hydrogels, without precipitating back out over time.

References

Bremmell, K. E. & Prestidge, C. A. (2019), 'Enhancing oral bioavailability of poorly soluble drugs with mesoporous silica based systems : opportunities and challenges', *Drug Development and Industrial Pharmacy* **45**(3), 349–358.
URL: *https://doi.org/10.1080/03639045.2018.1542709*

Chatterjee, S. & Hui, P. C.-L. (2018), 'Stimuli-Responsive Hydrogels: An Interdisciplinary Overview', *Hydrogels - Smart Materials for Biomedical Applications* .

Chen, L., Chen, B., Deng, L., Gao, B., Zhang, Y., Wu, C., Yu, N., Zhou, Q., Yao, J. & Chen, J. (2017), 'An optimized two-vial formulation lipid nanoemulsion of paclitaxel for targeted delivery to tumor', *International Journal of Pharmaceutics* **534**(1-2), 308–315.

Chuacharoen, T., Prasongsuk, S. & Sabliov, C. M. (2019), 'Effect of surfactant concentrations on physicochemical properties and functionality of curcumin nanoemulsions under conditions relevant to commercial utilization', *PubMed* **24**.

Cui, J., Yu, B., Zhao, Y., Zhu, W., Li, H., Lou, H. & Zhai, G. (2009), 'Enhancement of oral absorption of curcumin by self-microemulsifying drug delivery systems', *International Journal of Pharmaceutics* **371**(1-2), 148–155.

Eccleston, M. E., Kuiper, M., Gilchrist, F. M. & Slater, N. K. H. (2000), 'pH-responsive pseudo-peptides for cell membrane disruption', **69**, 297–307.

Evans, D. F., Pye, G., Bramley, R., Clark, A. G., Dyson, T. J. & Hardcastle, J. D. (1988), 'Measurement of gastrointestinal pH profiles in normal ambulant human subjects', **Gut, 1988**(29), 1035–1041.

Fahr, A., Liu, X., Fahr, A. & Liu, X. (2007), 'Drug delivery strategies for poorly water-soluble drugs Drug delivery strategies for poorly water-soluble drugs', *Expert Opinion on Drug Delivery* **4**(4), 403–416.

Farokhzad, O. C. & Langer, R. (2009), 'Impact of nanotechnology on drug delivery', *ACS Nano* **3**(1), 16–20.

Finosh, G. T. (2013), 'Influence of plasma protein-hydrogel interaction moderated by absorption of water on long-term cell viability in amphiliphilic biosynthetic hydrogels', *RSC Advances* (March).

Franck, A. (n.d.), 'Viscoelasticity and dynamic mechanical testing'.

Grumezescu, A. M. (2019), *Nanoarchitectonics in Biomedicine*, Elsevier.

Hampton, C., Demoin, D. & Glaser, R. E. (2010), 'Vibrational Spectroscopy Tutorial: Sulfur and Phosphorus'.

Höhne, G. W. H. (2019), 'Does anyone know why my DSC peak is overlapping?'.

Jabs, A. (n.d.), 'Determination of secondary structure in proteins by fourier transform infrared spectroscopy (ftir)'.
URL: *http://jenalib.leibniz-fli.de/ImgLibDoc/ftir/IMAGE_FTIR.html*

Jaiswal, M., Dudhe, R. & Sharma, P. (2014), 'Nanoemulsion: an advanced mode of drug delivery system', *Springer* **5**, 123–127.

Kelmann, R. G., Kuminek, G., Teixeira, H. F. & Koester, L. S. (2007), 'Carbamazepine parenteral nanoemulsions prepared by spontaneous emulsification process', *Elsevier Ltd.* **342**, 231–239.

Kim, C. K., Cho, Y. J. & Gao, Z. G. (2001), 'Preparation and evaluation of biphenyl dimethyl dicarboxylate microemulsions for oral delivery.', *Journal of controlled release : official journal of the Controlled Release Society* **70**(1-2), 149–55.
URL: *http://www.ncbi.nlm.nih.gov/pubmed/11166415*

Kim, J. E. & Park, Y. J. (2017), 'Paclitaxel-loaded hyaluronan solid nanoemulsions for enhanced treatment efficacy in ovarian cancer', *International Journal of Nanomedicine* **12**, 645–658.

Kumar, M., Ali, A., Kaldhone, P., Shirode, A. & Kadam, V. J. (2010), 'Available online through Report on pharmaceutical approaches to colon targeted drug delivery systems', **3**(3), 470–473.

Lavasanifar, A., Samuel, J. & Kwon, G. S. (2002), 'Poly(ethylene oxide)-block-poly(L-amino acid) micelles for drug delivery', *Advanced Drug Delivery Reviews* **54**(2), 169–190.

Lovelyn, C. & Attama, A. A. (2011), 'Current State of Nanoemulsions in Drug Delivery', *Journal of Biomaterials and Nanobiotechnology* **02**(05), 626–639.

Maitra, J. & Shukla, V. K. (2014), 'Cross-linking in Hydrogels - A Review', *American Journal of Polymer Science* **4**(4), 25–31.

Malvern Instruments Ltd. (n.d.), *Zeta Potential An Introduction in 30 Minutes*, Malvern Instruments Ltd.

Martin, L., Wilson, C. G., Koosha, F. & Uchegbu, I. F. (2003), 'Sustained buccal delivery of the hydrophobic drug denbufylline using physically cross-linked palmitoyl glycol chitosan hydrogels', *European Journal of Pharmaceutics and Biopharmaceutics* **55**, 35–45.

Nikam, T., Patil, M., Patil, S., Vadnere, G. & Lodhi, S. (2018), 'Nanoemulsion: A brief review on development and application in parenteral drug delivery', *Advance Pharmaceutical Journal* **3**, 43–54.

Philip, A. K. & Philip, B. (2010), 'Colon Targeted Drug Delivery Systems :', **25**(2), 70–78.

Rachmawati, H., Budiputra, D. K. & Mauludin, R. (2015), 'Curcumin nanoemulsion for transdermal application: Formulation and evaluation', *Drug Development and Industrial Pharmacy* **41**(4), 560–566.

Sant, V. P., Smith, D. & Leroux, J.-c. (2005), 'Enhancement of oral bioavailability of poorly water-soluble drugs by poly (ethylene glycol) - block -poly (alkyl acrylate- co -methacrylic acid) self-assemblies', *Journal of Controlled Release* **104**, 289–300.

Sari, T. P., Mann, B., Kumar, R., Singh, R. R., Sharma, R., Bhardwaj, M. & Athira, S. (2015), 'Preparation and characterization of nanoemulsion encapsulating curcumin', *Food Hydrocolloids* .

Shakeel, F., Baboota, S., Ahuja, A., Ali, J., Faisal, M. & Shafiq, S. (2008), Stability evaluation of celecoxib nanoemulsion containing Tween 80, Technical report.

Sharma, N., Mishra, S., Sharma, S., Deshpande, R. & Sharma, R. (2013), 'Preparation and optimization of nanoemulsions for targeting drug delivery', *International Journal of Drug Development and Research* **5**, 37–48.

Sood, S., Jain, K. & Gowthamarajan, K. (2014), 'Optimization of curcumin nanoemulsion for intranasal delivery using design of experiment and its toxicity assessment', *Colloids and Surfaces B: Biointerfaces* **113**, 330–337.

Stirrup, V., Ledingham, S. J., Thomas, M., Pye, G. & Evans, D. F. (1990), 'Redox potential measurement in the gastrointestinal tract in man', *Proceedings of the World Congress of Gastroenterology, Digestive Endoscopy and Colo-proctology* .

Wagner, J. G., Gerard, E. S. & Kaiser, D. G. (1966), 'The effect of the dosage form on serum levels of indoxole', *Clinical Pharmacology & Therapeutics* **7**(5), 610–619.
URL: *http://doi.wiley.com/10.1002/cpt196675610*

Wang, S. (2018), Synthesis and Characterisation of Stimuli-responsive Amino Acid-based Polymeric Materials for Drug Delivery, PhD thesis, Imperial College London.

Wang, S., Attah, R., Li, J., Chen, Y. & Chen, R. (2018), 'A pH-Responsive Amphiphilic Hydrogel Based on Pseudopeptides and Poly (ethylene glycol) for Oral Delivery of Hydrophobic Drugs', *ACS Biomaterials Science & Engineering* **4**, 4236–4243.

Wang, S., Ha, Y., Huang, X., Chin, B., Sim, W. & Chen, R. (2018), 'A New Strategy for Intestinal Drug Delivery via pH-Responsive and Membrane-Active Nanogels', *ACS Applied Materials & Interfaces* **10**, 36622–36627.

Wang, S., Liu, X., Villar-garcia, I. J. & Chen, R. (2016), 'Amino Acid Based Hydrogels with Dual Responsiveness for Oral Drug Delivery', pp. 1–7.

Wang, X., Jiang, Y., Wang, Y. W., Huang, M. T., Ho, C. T. & Huang, Q. (2008), 'Enhancing anti-inflammation activity of curcumin through O/W nanoemulsions', *Food Chemistry* **108**(2), 419–424.

Wang, Z. (2019), Development and characterization of stimuli-responsive amino acid-based hydrogel for drug delivery, Master thesis, Imperial College London.

Watkins, K. A. & Chen, R. (2015), 'pH-responsive , lysine-based hydrogels for the oral delivery of a wide size range of molecules', *International Journal of Pharmaceutics* **478**(2), 496–503.
URL: *http://dx.doi.org/10.1016/j.ijpharm.2014.12.005*

Wilding, I. R., Davis, S. S. & O'Hagen, D. T. (1994), 'Targeting of drugs and vaccines to the gut', *Pharmac. Ther.* **62**, 97–124.

Xu, W., Ling, P. & Zhang, T. (2013), 'Polymeric Micelles, a Promising Drug Delivery System to Enhance Bioavailability of Poorly Water-Soluble Drugs', *Journal of Drug Delivery* **2013**(1), 15.

Author Index

Linked to paper number

Adam, Muhammad 53
Adebajo, Esther 54
Agu, Ferdinand 40
Al-Riyami, Yousuf 54
Azimrayat Andrews, Thomas 21
Bade, Isha 40
Baldwin, Alexander 12
Behmer, Carl Fredrik 58
Bird, Max 25
Bird, Maxim S.F. 41
Chin, Teck Yang 39
Chitre, Aniket 3
Chaudhry, Arfan 1
Daniels, Toby 37
Eu Kane, Adrian Lua 8
Eu, Han Chang 8
Fan, Lanxin 48
Farmer, Matthew 19
Freake, Daniel 3
Fumagalli Romario, Giulio L. 9
García Gauna, Ane 45
Gocouladasse, Ramya 5
Hafiz, Usmaan 12
Han, Seok Ju 42
He, Xiaohan 7
Ho, Yuen Yin 4
Hoo, Swee Kee 33
Huang, Guan 32
Ismail, Iylia 33
Jaitly, Kathryn 26
Jiao, Yimeng 35
Juergensen, Tomas 21
Kamran, Athar 26
Kong, Weixuan 20
Kuek, Keneu 2
Lai, Yian-Shin 23
Lam, Derek 13
Latcham, Riley 24
Laws, Ben 15
Ledingham, Jack 60
Liang, Enshu 31
Lim, Kai L. 34
Liu, Menghan 35
Liu, Yifan 7
Lowe, Brandon 10
Mabonga, Amy 15
MacNeil, Jamie 22
Majekodunmi, Temi 28
Masud, Abrar 55
Mawdsley, Katie 50
Mehta, Komal 50
Mediavilla Madrigal, Sofía 45
Mehta, Shalin 29
Mitchell, Hamish 16
Morrissey, James 16
Motteu, Sophia 46
Ng, Yen Yiu 4
Nirmal, Yash 51
Nisar, Suhaib 18
O'Connor, Fintan 11
Ong, Titus 32
Ooi, Whai Shin 6
Patel, Huzaifa 51
Patel, Rhea 1
Patel, Vinay 14
Pillay, Bhanu P. 37
Potter, Jack 52
Ravindran, Abirami 59
Robert, Natsuko 44
Rovelli, Andrea 19
Sarkis, Miriam 44
Satura, Tomas 46
Savulescu, George-Claudiu 30
Shaieb, Remy 27
Shah, Khilan 5
Shah, Roshan 43
Shen, Richard 39
Sherbaz, Hassan 57
Sifneos, Maximos 14
Song, Zijie 43
Tam, Klement 57
Tan, Cheow Lin 23
Tangtrakulwong, Pichan 6
Teck, Anastasia A. 38
Teo, Tricia W. T. 34
Thapelo, Nonofo 17
Tsolope, Tumisang 17
Tung, Su Cheng 22
Ueda, Maiki 42
van de Berg, Damien 58
Vasey, Alexa 13
Waabis, Muhammad 55

Wang, Xiao Yuan 10
Ward, Adam 25
Watson, Abbie 27
Wild, India C 9
Wong, Chloe 59
Wong, Ian 52
Wu, Yue 31
Wulkau, Yasmin 28
Xie, Mingrou 47
Xiong, Haoran 53
Xu, Bohan 41
Yap, Fui Ling 49
Yeung, Christopher 11
Yew, Hon Wa 47
Yu, Luxi 56
Yu, Zhaomin 20
Zangeneh, Saurav 29
Zhang, Shengming 38
Zhang, Yu 36
Zhou, Xiaoqun 56

Supervisor Index

Linked to paper number

Adjiman, Claire S.	36, 44
Campbell, Kyra	60
Chachuat, Benoit	32, 41
Chadwick, David	15, 28
Chen, Rongjun	7, 31, 59
Fennell, Paul	1, 17
Hallett, Jason	48
Hawkes, Adam	22, 54, 57
Heng, Jerry	6, 40
Hellgardt, Klaus	4, 50
Kazarian, Sergei	10, 23
Krishnan, J	29
Li, Kang	20, 42, 53
Livingston, Andrew	21
Luckham, Paul	30, 33, 51
Markides, Christos	46
Matar, Omar K.	2, 47, 55
Müller, Erich A.	12, 25
Petit, Camille	11, 27, 43
Pini, Ronny	19, 34
Polizzi, Karen	13
Rinaldi, Roberto	52
Rücker, Maja	51
Shah, Nilay	16, 37, 49, 58
Song, Qilei	39, 56
Tighe, Chris	5, 8, 14, 18
Titirici, Magda	3, 9, 26, 38
Trusler, J.P. Martin	24
Williams, Daryl	45
Yetisen, Ali	35