

IMPERIAL

Towards Process Automation in Upstream Biopharmaceutical Processing

by Thomas Wayne Kavanagh

The Sargent Centre for Process Systems Engineering

Department of Chemical Engineering, Imperial College London

A thesis submitted to Imperial College London for the degree of

Doctor of Philosophy

2024

To my family including friends, and all others who have helped me get to where I am.

To all the barbeques and attendees who have helped me address the burning questions herein.

Declaration of Originality

I hereby certify that all material in this thesis that is not my own work has been appropriately acknowledged.

by Thomas Wayne Kavanagh

2024

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non-Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. 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. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. 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.

Acknowledgements

First and foremost, I extend my deepest gratitude to my supervisors, **Cleo Kontoravdi** and **Maria M. Papathanasiou**, the ultimate academic duo whose expertise and patience have been invaluable throughout my PhD journey. It has been a true privilege to work under your guidance. **Cleo**, you have shown me the power and elegance of simplicity. **Maria**, you have demonstrated how to deliver even the most technical concepts with energy and impact. Together, your unwavering support has consistently challenged me to think critically, question everything, and continuously grow. I am profoundly grateful to those at GSK who generously supported me throughout my PhD journey. **Paloma Diaz Fernandez**, **Pavlos Kotidis**, and **Gary Finka** thank you for facilitating valuable connections across GSK and helping to ensure my time at GSK had minimum disturbance. **Michele Albanese** and **Katy Newell**, I deeply appreciate your guidance through the complexities of an industrial process development laboratory. **Yu Luo** and **Sameer Talwar**, thank you for being consistently available to offer critique and thus help me to refine my ideas. In addition, thank you to **Gabriele Bano**, **Zachary Hatzenbeller**, and **Anastasia Nikolakopoulou** for warmly welcoming me to the world-leading MPC Group at GSK. I want to extend my acknowledgements to **Nilay Shah** and **Alexandros Kiparissides** for offering your valuable insights at the end of my PhD journey. Next thank you **Imperial College London** and the **Department of Chemical Engineering** for providing an inspiring environment that made this research possible. The support from the **Biotechnology and Biological Sciences Research Council** and **GSK** is also gratefully acknowledged.

To my two academic families the Kontoravdi Group and the Papathanasiou Group, I am immensely appreciative of your motivation through challenging times and for the camaraderie that made this journey unforgettable over our breakfasts, lunches, and many, many coffees. Special thanks to Pavlos for being an outstanding role model; Sakhr for your unmatched hospitality; Athanasios, Rødrigo, James, Mariana and Ben, for bringing your outstanding humour to the office; Roberto, for your simple yet exceptional culinary ability; Simon, for our many adventures; and Miriam, for going above and beyond to unite us all. To my other immediate academic family, thank you for your contributions to this shared journey, including Elli, Ignacio, Zoltan, Masue, Chiara, Steven, Haiting, Konstantinos, Niki, Foteini, Syazana,

Luxi, Jacopo, Dilyara, Andrea, and Daniele. To my academic cousin Tanuj, our friendship has been a cornerstone of my time at Imperial. Together, our shared passion for bringing people together has made this journey more meaningful. To others including, Ethan, Saher, David, Kam, Dion, Ahmed, Akhil, Alessandra, Tom, Shubhani, Nikki, Deniz, Oscar, Nadine, Sean, Mathias, Niki, Ilya, Gustavo, Lucy, Celine, Michele, Zhengang, and Valentina thank you for your companionship and for bringing life to GEF and musical talent to the Control Club.

Additionally, thanks to Jarka, Pier, Paul, Milan, Chris and Rick for your support and encouragement towards pursuing a PhD. A heartfelt thanks to St George's Group including Kirsty, Eloise, Natalie, Max, and Ellen. Here is to many more reunions. Further, thank you to the Oxford Group including Neftalem, Hanna, Codie, Will, Katie, Angel, Tolu, and Zonerah, soon we will punt again. I extend additional acknowledgement to the NCL Group. Lucien and Hayden, you've both been family to me for many years. Life would be far less fun without the many BBQs, cheffing, furniture building, adventures, and events we have organised together. Ulas and Osman the exquisite culinary experiences we have shared over the years have been a source of strength. Osei, your positivity is infectious, and I am deeply grateful for your steadfast support. Luke, we all miss you, thank you for being there throughout the years, it was a pleasure. Sam, it was a joy spending the first year and lockdown with you. Also, thanks to the following who have kept me going in various ways but whose contributions would make this section longer than my thesis including Brad, Anis, Ahmed, Deej, Will, Anujan, Anna, Lottie, Santiago, Ash, Ajax, Munira, Vicky, Anam, Lise, Eleni, Tasos, Natasha, Louise, and especially Ella for your incredible support. Thank you to Rees, Fred again.., Bicep, Josh FB, Chris Luno, Krept & Konan, and many more whose support has been music to my ears. Finally, thank you to the Boro Group who were there throughout my PhD offering much-needed parmos especially Andy and Jack for picking me up when times were tough, and James for giving me the best sendoff. In addition, thanks to Ben, Ali, Owen, Lewis, Adam, Perry, Joe, Payten, Charlie, Declan, and Matty.

Last but certainly not least, thank you to my family especially all those who played a part in helping me get to where I am. I would particularly like to thank my Parents for your amazing support. Finally, to Fran, your humour, and even attempted humour, never fails to raise spirits.

To everyone else who contributed to this journey, I am forever grateful.

Abstract

Monoclonal antibodies (mAbs) continue to show significant clinical and commercial success in treating major diseases including heart diseases, cancers, and Alzheimer's. However, the great costs of biologics, and the development of biosimilars, have driven demand for high-performing and lower-cost manufacturing methods. Process automation can unlock such benefits, but its application to biopharmaceutical processes is impeded by challenges including a lack of inexpensive sensors and difficulty controlling the complex process dynamics. This thesis employs process systems engineering to overcome these barriers advancing towards process automation of biopharmaceutical processes. Firstly, modelling workflows are demonstrated capable of elucidating the behaviour of highly complex processes in instances of limited and expensive data. From this, one can gain an in-depth understanding whilst alleviating the experimental burden of process development. To that end, several experiments were performed characterizing a commercial biopharmaceutical cell culture process under varying feeds of glucose. Next, a mechanistic mathematical model designed for real-time deployment was postulated to describe Chinese hamster ovary (CHO) cell growth and metabolism and subsequently applied to devise an optimal operating strategy. Finally, multiple novel process control strategies were developed and demonstrated experimentally both during a major process deviation and during normal operation. Namely, in an attempt to recover a batch during a major process deviation, reactive scheduling and non-linear model predictive control (NLMPC) were deployed. Further, during normal operation, open-loop optimisation was compared to NLMPC. The control objective of all demonstrations was identical and included maximising product titre whilst minimising exposure to lactate, a metabolic by-product that adversely affects cell culture performance. Here, NLMPC satisfied the control objective achieving superior performance. Overall, this research aims to pave the way towards more advanced process control and automation of biopharmaceutical production processes with the overarching goals of increasing process performance and robustness, reducing process development timelines, and improving process economics.

Contents

Contents	xii
List of Figures	xv
List of Tables	xix
Nomenclature	xxi
Chapter 1 Introduction.....	1
Chapter 2 Background & Literature Review	6
2.1 Overview of Monoclonal Antibodies.....	6
2.2 Production of Monoclonal Antibodies.....	9
2.3 Biopharmaceutical Process Development.....	10
2.4 Process Measurement Technologies for Cell Cultures	12
2.5 Influence of Conditions on CHO Cell Cultures	15
2.5.1 Impact of Temperature.....	15
2.5.2 Impact of pH	16
2.5.3 Impact of Dissolved Oxygen	17
2.5.4 Impact of Feed	17
2.6 Modelling Mammalian Cell Cultures	19
2.6.1 Model Classification	19
2.6.2 Recent Advances in Modelling CHO Cell Cultures	25
2.6.3 Challenges in Modelling Cell Cultures.....	32
2.7 Overview of Control of Mammalian Cell Cultures	32
2.7.1 Control Problem Definition	33
2.7.2 Control Algorithm Classification.....	33
Chapter 3 Project Aims.....	50
3.1 Research Proposal.....	50
3.2 Research Objectives & Proposal.....	51
Chapter 4 Mechanistic Modelling of CHO Cell Cultures for Online Application.....	56
4.1 Introduction.....	56
4.2 Materials & Methods	59
4.2.1 Cell Culture.....	59
4.2.2 Analytical Methods.....	60
4.2.3 Data Processing and Analysis.....	61

4.2.4	Mathematical Modelling.....	65
4.2.5	Data Normalization.....	67
4.3	Experimental Observations.....	68
4.4	Model Development.....	72
4.5	Model Parameterisation	83
4.6	Results.....	85
4.6.1	Model Parameter Sensitivity Analysis.....	85
4.7	Model Performance.....	87
4.8	Discussion	90
4.9	Summary	93
Chapter 5	Advanced Process Control of CHO Cell Cultures.....	94
5.1	Introduction.....	94
5.1.1	Operating Strategy Development.....	95
5.1.2	Open-Loop vs Closed-Loop Systems	96
5.1.3	Key Process Control Algorithms	96
5.1.4	Developments in Model Predictive Control of CHO Cell Cultures	98
5.1.5	Summary	104
5.2	Materials & Methods	106
5.2.1	Cell Culture.....	106
5.2.2	Analytical Methods.....	107
5.2.3	Mathematical Modelling.....	107
5.2.4	Process Control Implementation.....	109
5.2.5	Operational Strategy Development.....	113
5.3	Operational Strategy Design.....	115
5.3.1	Input-Output Analysis.....	116
5.3.2	Input-Output Sensitivity Analysis.....	123
5.3.3	<i>In silico</i> Optimisation.....	129
5.4	Results.....	134
5.4.1	Variation in Manipulated Variable	134
5.4.2	Reactive Scheduling vs NLMPC During an Operational Error.....	136
5.4.3	Open-loop Optimisation vs NLMPC during Normal Operation.....	139
5.4.4	Experimental Observations vs In-Silico Optimisation	142
5.4.5	Comparison of Variables' Endpoint Value.....	143
5.5	Discussion.....	145
5.5.1	Comparing Variations in Manipulated Variables	146
5.5.2	Comparing Reactive Scheduling and NLMPC during an Operational Error	147

5.5.3	Comparing Open-loop Optimisation and NLMPC during Normal Operation ..	149
5.5.4	Comparison of In-Silico Optimisation vs Experimental Observations.....	151
5.5.5	Comparison of all Process Control Strategies	152
5.6	Summary	154
Chapter 6	Conclusions and Further Research.....	157
6.1	Main Contributions	159
6.2	Limitations & Future Research.....	162
6.2.1	Improvements to the Mechanistic Model	162
6.2.2	Improvements to the Process Controller.....	166
6.3	Concluding Remarks.....	168
Chapter 7	References.....	169

List of Figures

Figure 1 Structure of an IgG antibody.	6
Figure 2 Schematic of a biopharmaceutical process.	9
Figure 3 Process development workflow.	10
Figure 4 Mathematical model classifications. Adapted from (Tangirala, 2018).	19
Figure 5 Scales across a cell culture.	22
Figure 6 Open-loop system.	33
Figure 7 Closed-loop control.	37
Figure 8 Process-integral-derivative control.	38
Figure 9 Model predictive control	41
Figure 10 Towards implementation of model predictive control to fed-batch production of monoclonal antibodies.	52
Figure 11 Model-building experiment overview with selected key outputs listed. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose.	60
Figure 12 Normalised volume of glucose fed on each day throughout the model-building experiments. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose. Volumes are normalised based on the maximum volume of glucose fed which occurred in MF-H on Day 10.	60
Figure 13 Automated data processing - workflow built in Python using NumPy.	62
Figure 14 Automated data processing-- unit conversion and variable labelling workflow performed in Python.	63
Figure 15 Mechanistic modelling workflow.	65
Figure 16 Effects on cell culture of a nominal glucose feeding strategy (MF-N), a low glucose feeding strategy (MF-L), and a high glucose feeding strategy (MF-H). Normalised using min-max normalisation.	70
Figure 17 Specific productivity across time within the experimental model-building experiments.	71
Figure 18 Specific growth rate across time within the experimental model-building experiments.	71

Figure 19 Specific death rate across time within the experimental model-building experiments. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose. Normalised using min-max normalisation.....	71
Figure 20 Viable cell density across time within the experimental model-building experiments: MF-N is with nominal glucose feeding, MF-L is with low glucose feeding, and MF-H is with high glucose feeding. Segmentation based on cell growth stage is colour-coded. Normalised using min-max normalisation.	74
Figure 21 Specific growth rate across time within the experimental model-building experiments: MF-N is with nominal glucose feeding, MF-L is with low glucose feeding, and MF-H is with high glucose feeding. Segmentation based on cell growth stage is colour-coded. Normalised using min-max normalisation.....	74
Figure 22 Relationship of growth and metabolite utilisation in (a) major metabolites in MF-N, (b) minor metabolites in MF-N, (c) major metabolites in MF-L, (d) minor metabolites in MF-L, (e) major metabolites in MF-H, (f) minor metabolites in MF-H. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose. Normalised using min-max.	78
Figure 23 Effects of osmolality on specific growth rate in model-building experiments with nominal glucose feeding (MF-N), low glucose feeding (MF-L), and high glucose feeding (MF-H). Normalised using min-max normalisation.....	79
Figure 24 Death rate across the cell culture period. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose. Normalised using min-max normalisation.	80
Figure 25 First order indices and total sensitivity indices from the parameter sensitivity analysis varying parameters by 40% on (a) summing indices on each parameter, and (b) summing indices on each output.....	86
Figure 26 Experimental vs model predictions of variables within the mechanistic model across the cell culture period, and in parity. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose. Normalised using min-max normalisation.....	89
Figure 27 Open-loop optimisation architecture used for the OLO-N experiment. OLO-N: open-loop optimisation experiment in the absence of major operating errors.....	111
Figure 28 Reactive scheduling architecture used for the RS-E experiment. OF: Objective function. IGs: Initial guesses. <i>td</i> : time of major process deviation. RS-E: reactive scheduling in the presence of a major operating error.	112
Figure 29 Model predictive control architecture used for experiments MPC-E and MPC-N. OF: Objective function. IGs: Initial guesses. MPC-E: model predictive control in the presence of a major operating error. MPC-N: model predictive control in the absence of major operating errors.	113
Figure 30 Input-output analysis comparing, at each day, the total feed volume of glucose to the total cell density and product titre. Normalised using min-max normalisation.....	116
Figure 31 Input-output analysis comparing, at each day, the total feed volume of glucose to the viable cell density and product titre. Normalised using min-max normalisation.	117

Figure 32 Input-output analysis comparing, at each day, the total feed volume of glucose to the extracellular concentration of glucose and product titre. Normalised using min-max normalisation.....	119
Figure 33 Input-output analysis comparing, at each day, the total feed volume of glucose to the extracellular concentration of lactate and product titre. Normalised using min-max normalisation.....	121
Figure 34 Input-output analysis comparing, at each day, the total feed volume of glucose to the extracellular osmolality and product titre. Normalised using min-max normalisation.	122
Figure 35 First order sensitivity indices showcasing the effects of each scheduled glucose feed on the process outputs at the sample time after feeding, and prior to the subsequent feed. Normalised using min-max normalisation.....	125
Figure 36 Total effects indices showcasing the effects of each scheduled glucose feed on the process outputs at the sample time after feeding, and prior to the subsequent feed. Normalised using min-max normalisation.	126
Figure 37 Ratio between the first order sensitivity indices and total effects indices showcasing the effects of each scheduled glucose feed, and the value of the process outputs at the sample time after feeding, and prior to the next feed. Normalised using min-max normalisation. ...	127
Figure 38 Results from <i>in silico</i> optimisation. Normalised using min-max normalisation based on the max of each endpoint value.	130
Figure 39 Optimal feeding schedule and <i>in silico</i> output trajectories for S6. Normalised using min-max normalisation based on a maximum value on the y-axis.....	132
Figure 40 Relationship between manipulated variables, objective function, and product titre. Normalised using min-max normalisation.....	133
Figure 41 Variation in daily glucose feed due to the process control strategy. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a high amount of glucose. RS-E: reactive scheduling in the presence of a major operating error. MPC-E: model predictive control in the presence of a major operating error. OLO-N: open-loop optimisation experiment in the absence of major operating errors. MPC-N: model predictive control in the absence of major operating errors. Normalised using min-max normalisation based on the overall maximum volume of glucose fed.....	134
Figure 42 Variation in total glucose feed due to the process control strategy. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a high amount of glucose. RS-E: reactive scheduling in the presence of a major operating error. MPC-E: model predictive control in the presence of a major operating error. OLO-N: open-loop optimisation experiment in the absence of major operating errors. MPC-N: model predictive control in the absence of major operating errors. Normalised using min-max normalisation in Figure 41 then totalled in this Figure.....	135
Figure 43 Trajectories of model predictions versus experimental observations when deploying reactive scheduling (RS-E) and non-linear model predictive control (MPC-E) during an operational error. Normalised using min-max normalisation.	137
Figure 44 Absolute percentage error between model predictions versus experimental observations when deploying reactive scheduling (RS-E) and non-linear model predictive control (MPC-E) during an operational error.	138

Figure 45 Trajectories of model predictions versus experimental observations when deploying open-loop optimisation (OLO-N) versus non-linear model predictive control (MPC-N) during normal operating conditions. Normalised using min-max normalisation.....	140
Figure 46 Absolute percentage error between model predictions versus experimental observations when deploying open-loop optimisation (OLO-N) and non-linear model predictive control (MPC-N) during normal operating conditions. Normalised using min-max normalisation.....	141
Figure 47 Percentage error between the anticipated optimal performance and achieved performance of variables within the objective function. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a high amount of glucose. RS-E: reactive scheduling in the presence of a major operating error. MPC-E: model predictive control in the presence of a major operating error. OLO-N: open-loop optimisation experiment in the absence of major operating errors. MPC-N: model predictive control in the absence of major operating errors.	143
Figure 48 Comparison of all endpoint values of variables measured during cell culture experiments normalised against the manual feeding experiment with the highest product titre, MF-L.	144

List of Tables

Table 1 Example process control strategy defined using QbD and adapted from (J. Xu et al., 2022)	12
Table 2 Overview of key process parameters monitored by PAT during fed-batch CHO cell culturing, adapted from (Reyes et al., 2022).....	14
Table 3 Recent advances in modelling CHO cell cultures. For research listed here, a literature review network has been built using Inciteful (follow link) which lists an up-to-date version of similar papers, review papers, predicts top authors, and predicts important papers (Weishuhn, 2024a).	25
Table 4 Examples of open-loop systems deployed for mammalian cell culturing. All defined using model-based optimisation.	35
Table 5 Examples of closed-loop control deployed for mammalian cell culturing.	40
Table 6 Examples of model predictive control deployed for mammalian cell culturing.....	42
Table 7 Recent advances in control of CHO cell cultures. For research listed here, a literature review network has been built using Inciteful (follow link) which lists similar papers, review papers, predicts top authors, and predicts important papers (Weishuhn, 2024b).	45
Table 8 Summary of methods utilised within this thesis.	53
Table 9 Automated data processing— unit conversion example moving from raw data to standardised data.....	63
Table 10 Quantitative metrics defining the relationship between growth rate and metabolite utilisation rates of minor metabolites.....	76
Table 11 Quantitative metrics defining the relationship between growth rate and metabolite utilisation rates of major metabolites.	76
Table 12 Parametrisation results for the parameters associated with growth and productivity in the mechanistic model. *95% confidence interval is greater than 10% of the parameter value.	84
Table 13 Parametrisation results for the parameters associated with extracellular metabolite concentration and osmolality in the mechanistic model.	84
Table 14 Performance metrics for the goodness of fit between model predictions and experimental data. RMSE: Root mean square error, PCC: Pearson Correlation Coefficient. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose.....	89
Table 15 Summary of key contributions implementing model predictive control to CHO cell cultures. MVs: manipulated variables. CVs: controlled variable. OF: objective function. FB:	

fed-batch. DD: data-driven. SPEM: setpoint error minimisation. CVM: CV minimisation/maximisation.....	99
Table 16 <i>In silico</i> optimisation strategies. +: Maximise. -: Minimise.	129
Table 17 Variance and standard deviation of the volume of glucose fed across process control strategies.	135
Table 18 Model performance when deploying reactive scheduling (RS-E) and non-linear model predictive control (MPC-E) during an operational error. RMSE: root mean square error in units of measurements for X_t (cells.mL ⁻¹), X_v (cells.mL ⁻¹), [mAb] (mg.L ⁻¹), [Glc] (mM), [Lac] (mM), Osm (mOsmol.Kg ⁻¹). PCC: Pearson correlation coefficient.	139
Table 19 Model performance when deploying open-loop optimisation (OLO-N) and non-linear model predictive control (MPC-N) during normal operating conditions. RMSE: root mean square error in units of measurements for X_t (cells.mL ⁻¹), X_v (cells.mL ⁻¹), [mAb] (mg.L ⁻¹), [Glc] (mM), [Lac] (mM), Osm (mOsmol.Kg ⁻¹). PCC: Pearson correlation coefficient.	142

Nomenclature

K_d	Derivative gain
K_I	Integral gain
K_p	Proportional gain
$C_{i,feed}$	Concentration of metabolite i in the feed (mM)
C_i	Extracellular concentration of metabolite i (mM)
F_{in}	Volumetric flow rate in (L.h ⁻¹)
F_{out}	Volumetric flow rate out (L.h ⁻¹)
K_{Lysis}	Rate of cell lysis (h ⁻¹)
K_{OsmLac}	Kinetic term for lactate utilisation (mOsm/kg H ₂ O)
K_i	Limiting constant
X_d	Dead cell density (cells.L ⁻¹)
X_t	Total cell density (cells.L ⁻¹)
X_v	Viable cell density (cells.L ⁻¹)
Y_{LacGlc}	Yield of lactate from glucose (mM _{Lac} .mM _{Glc})
f_{inh}	Inhibitory factor (-)
f_{lim}	Limiting factor (-)
$m_{Lac,1}$	Lactate maintenance coefficient (mM.cell ⁻¹ .hr ⁻¹)
$m_{Lac,2}$	Kinetic term for lactate utilisation (mM.cell ⁻¹ .hr ⁻¹)
m_{mAb}	Non-growth associated term for productivity

q_i	Cell-specific metabolite utilisation rate (mM.cell ⁻¹ .h ⁻¹)
t_0	Initial time
t_f	Final time
$u_{SP}(t)$	Setpoint trajectory
α_{Lac}	Kinetic term for lactate utilisation (mOsm/kg H ₂ O).
$\alpha_{fin,osm}$	Kinetic term for effect of feed on osmolality (mOsm/kg H ₂ O/L/h)
$\alpha_{glc,osm}$	Kinetic term for the effect of glucose on osmolality (mOsm/kg H ₂ O/mM)
$\alpha_{lac,osm}$	Kinetic term for the effect of lactate on osmolality (mOsm/kg H ₂ O/mM)
α_{osm}	Kinetic term for osmolality (mOsm/kg H ₂ O)
μ_{death}	Cell-specific death rate (h ⁻¹)
$\mu_{growth,max}$	Maximum cell-specific growth rate (h ⁻¹)
$\mu_{growth,overall}$	Overall growth rate (h ⁻¹)
μ_{growth}	Cell-specific growth rate (h ⁻¹)
[Met]	Concentration of metabolite (mM) or monoclonal antibody (mg.L ⁻¹)
Amm	Ammonia
ATP	Adenosine-triphosphate
CHO	Chinese hamster ovary
CMA	Critical material attributes
CPP	Critical process parameters
CQA	Critical quality attributes
DO	Dissolved oxygen
DoDE	Design of dynamic experiments
DOT	Dissolved oxygen tension
EMA	European Medicines Agency
FDA	US Food and Drug Administration

Glc	Glucose
Gln	Glutamine
Glu	Glutamate
HCP	Host cell protein
IntFin	Total feed in (L)
IntLac	Integral lactate concentration (mM)
IntOsm	Integral osmolality (mOsm/kg H ₂ O)
IVCD	Integral viable cell density (cells.L ⁻¹)
KPI	Key performance indicators
KPP	Key process parameters
Lac	Lactate
LDH	Lactate dehydrogenase
mAb	Monoclonal antibody
MPC	Model predictive control
NIR	Near-infrared
Osm	Osmolality (mOsm/kg H ₂ O)
OUR	Oxygen uptake rate
PAT	Process analytical technology
PCC	Pearson correlation coefficient
PE	Percentage error
PID	Proportional-integral-derivative control
QbD	Quality by design
QbT	Quality by testing
RMSE	Root mean square error
Si	First order sensitivity index
SP	Setpoint

St	Total order sensitivity index
t	Time (h)
T	Time horizon
TCA	Tricarboxylic acid
tcOUR	Cumulative oxygen uptake rate
u(t)	Controller input at time t
V	Working volume of the cell culture (L)
WMPE	Weighted mean percentage error
KI_i	Inhibitory constant
Lac_{max}	Kinetic constant for lactate consumption (mM)
$x(t)$	Differential variable
$y(t)$	Algebraic variable
μ	Mean
σ	Standard deviation
α	Gradient from a linear regression between the cell-specific growth rate and the Cell-specific metabolite utilisation rate
MF-N	Experiment abbreviation: manual feeding a nominal amount of glucose
MF-L	Experiment abbreviation: manual feeding a low amount of glucose
MF-H	Experiment abbreviation: manual feeding a high amount of glucose
RS-E	Experiment abbreviation: reactive scheduling in the presence of a major operating error
MPC-E	Experiment abbreviation: model predictive control in the presence of a major operating error
OLO-N	Experiment abbreviation: open loop optimisation in the absence of major operating errors.
MPC-N	Experiment abbreviation: model predictive control in the absence of major operating errors.

Chapter 1 Introduction

The biopharmaceutical industry has a major positive impact on modern society, making progress towards the United Nations Sustainable Development Goal 3 to “ensure healthy lives and promote well-being for all at all ages” (United Nations, 2023). Biopharmaceuticals are derived from biological sources and represent some of the greatest accomplishments in modern science (Kesik-Brodacka, 2018). Decades of research and development have led to treatments continuing to emerge for many of the leading causes of death globally (in 2023) including heart diseases (e.g., evolocumab), cancers (e.g., trastuzumab deruxtecan), and Alzheimer’s (e.g., lecanemab) (B. T. Li et al., 2022; Sabatine et al., 2017; Swanson et al., 2021; World Health Organization, 2020). In 2021, the biopharmaceuticals market was substantial with global sales estimated to be worth \$343 billion (G. Walsh & Walsh, 2022) equivalent to the gross domestic product of more than three-quarters of the countries and economies ranked by the World Bank (World Bank & OECD, 2022). Although there is a growing interest in other modalities (e.g., cell and gene therapies), monoclonal antibodies (mAb) remain the dominant product, representing more than 80% of the global sales of biopharmaceuticals in 2021 (G. Walsh & Walsh, 2022). Mammalian cell systems are commonly used to produce mAb products, reflecting their ability for complex post-translational modifications such as glycosylation (G. Walsh & Walsh, 2022). Further, Chinese hamster ovary (CHO) cells are used in the production of 89% of approved products made in mammalian systems owing to decades of research and development that has produced highly productive cell lines which at present, are capable of delivering titres around 10 g.L^{-1} (Lalonde & Durocher, 2017; G. Walsh & Walsh, 2022). Finally, although there is increased adoption of continuous bioprocessing across the biopharmaceutical industry, fed-batch production represents the most common manufacturing method for the expression of monoclonal antibodies (Rish et al., 2022; Yee et al., 2018).

Despite the success of the biopharmaceutical industry, the adoption of modern Industry 4.0 technologies towards enabling process automation has lagged compared to other process

industries (MacDonald, 2020; Silva et al., 2020a). Across other process industries, such technologies have led to increased throughput, product quality improvements, downtime reductions, and lowering of operating costs by up to 15% (Bauer & Craig, 2008a; William Y. Svrcek, Donald P. Mahoney, 2014). However, within the biopharmaceutical industry, drug discovery is a key focus before process development (von Stosch et al., 2021). Even so, process development can increase the total time to market of biopharmaceuticals by 2-7 years (von Stosch et al., 2021). As such, there are many pressures on the biopharmaceutical industry to improve process development timelines and thus decrease the time from the drug discovery to the patient administration of biopharmaceuticals (Kelley, 2020; Pham et al., 2023). A recent success of this was demonstrated during the widely celebrated rapid timeline of the SARS-CoV-2 vaccine (Ball, 2021; Krammer, 2020). Aside from this, upcoming patent cliffs have also catalysed manufacturers to investigate methods of cost reduction (Ratih et al., 2021; Reddy et al., 2023; G. Walsh & Walsh, 2022). Notably, adalimumab-atto (Amgen) recently entered the market competing directly with one of the all-time best-selling biopharmaceuticals, adalimumab (AbbVie) (Amgen, 2023). The positive effects of reducing time and cost of process development should not be understated. For example, being only a quarter of a year faster to market is estimated to increase the net present value of drugs by 5% (Glass et al., 2016). In general, utilising Industry 4.0 enabling technologies to improve process development timelines may also lead to increased productivity, greater competitiveness, higher product quality, reduced waste, and lower operating costs (Silva et al., 2020b). Further, recent advances in drug discovery (e.g., using artificial intelligence) necessitate improved process development capabilities due to a predicted increased number of candidates whereby the limited parallelization ability of current process development workflows may create further bottlenecks in delivering treatments to patients (von Stosch et al., 2021). Overall, there is a clear value in focusing on process development of biopharmaceutical processes and here Industry 4.0 enabling technologies are expected to contribute towards the end vision of fully automated processes. With this, the biopharmaceutical industry can radically advance process development capabilities, directly addressing the United Nations Sustainable Development Goals 9 (build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation) and 12 (ensure sustainable consumption and production patterns) (United Nations, 2023).

Notwithstanding the above promises, progressing biopharmaceutical processing towards automation presents significant challenges and requires major advances in process modelling and process control. In line with this, regulatory agencies are enforcing stricter standards to ensure consistent product safety and efficacy including using the Quality by Design framework (Anurag S Rathore & Helen Winkle, 2009). Such guidelines include a requirement for a process knowledge base to understand and control processes more effectively. Gaining in-depth process understanding to design suitable operating strategies using solely experimental efforts is not only challenging but highly expensive and laborious. As such, computational approaches are becoming increasingly popular whereby a low number of experiments are used to produce a mathematical approximation of the process before subsequent computational studies are used to increase process understanding or develop operating strategies (Schmidt et al., 2024). Broadly, implementing in-silico modelling tools within process development can significantly reduce the length of process development timelines (Zobel-Roos et al., 2019). Meanwhile, further reductions in process cost, improvements in process performance, and increased process robustness may be achieved using advanced process control (Luo, Kurian, et al., 2023; Mitra & Murthy, 2022). Inherently this requires the capability to measure relevant process variables ideally online or at-line. To that end, the Process Analytical Technology (PAT) framework developed in 2004 by the Food & Drug Administration has led to much research (e.g. in Raman spectroscopy) hence improved sensor technology (Simon et al., 2015). (Biechele et al., 2015). That said, this does not mean that this technology is ready for, or widely implemented online, especially during large-scale manufacturing (Martínez-Monge et al., 2022). Therefore, the demonstration of a control strategy that leverages only existing widely deployed sensor technologies would be of value albeit challenging without state-of-the-art techniques.

Across research, there has been a strong focal point in modelling of upstream biopharmaceutical processes to improve process understanding or define operating strategies using mechanistic and data-driven modelling strategies (Kiparissides et al., 2011; Kyriakopoulos et al., 2018) with more recent models built for fed-batch CHO cell cultures to capture effects of osmolality and help reveal the lactate metabolism (Alhuthali & Kontoravdi, 2022; Luginsland et al., 2024). Whilst data-driven models have been deployed for real-time application frequently due to relative ease of development (Dewasme et al., 2023), the identification of mechanistic models including more rigorous parameterisation and structuring efforts required can be relatively burdensome (Doyle et al., 2023). Nevertheless, mechanistic

models have also been shown to outperform other types of models during instances of low data availability (Bayer et al., 2023). As such, it would be of merit to overcome the challenges and showcase the usefulness of mechanistic models built from limited data during deployment in real-time.

With regard to process control of CHO cell cultures, there is also limited research using advanced process control of key performance indicators. Yet there is a clear need for advanced control capable of tackling the strong non-linearities and time-varying dynamics present in such processes (Luo, Kurian, et al., 2021). Despite this, biopharmaceutical processes often rely on regulatory control using classical control techniques (e.g., proportional-integral-derivative controllers) within single input single output control loops (Rathore et al., 2021). Here, the ability to exert influence on the process is limited. Further, batch-to-batch variability can be exacerbated by highly coupled process inputs where for example, manipulating one process input risks destabilising a control variable in a separate control loop. However, the development of an advanced process controller is challenging and requires a suitable process model to be built prior to testing and deployment. In addition, the costs associated with generating experimental data for modelling from upstream biopharmaceutical processes are great. Here lies an opportunity for the deployment of a mechanistic model within a model-based controller for multi-output closed-loop control of key performance indicators.

In this thesis, several computational tools are harnessed and demonstrated to advance the production of monoclonal antibodies towards process automation. Briefly, the cell culture is modelled, and advanced process control is applied. Firstly, background is given to biopharmaceuticals, biopharmaceutical processes, as well as the current status quo in biopharmaceutical manufacturing (Chapter 2). Next, the state of the art in modelling and control of CHO cell cultures is presented setting the basis for, and demonstrating the importance of, the work within this thesis (Chapter 2).

An overview of the materials and methods is given in Chapter 3. Further, detailed descriptions are integrated into each results chapter separately to enable an increased understanding of each study. Initially, experiments are performed in a single-use bioreactor using an industrial cell line previously used at a manufacturing scale (Chapter 4). The process is interrogated by varying a glucose feeding schedule to the cell culture and analysing the responses of key process variables including product titre, viable cell density and metabolites. With the aim of

moving towards automation, a generalisable data analysis script is built taking raw experimental data and outputting various important metrics e.g., growth rates. After data analysis, modelling logic is demonstrated that selects out of the measured metabolites, which are most influential on growth. Once the model was built, it was tested against experimental data to assess performance. Overall, a limited number of experiments were performed allowing the power of mechanistic models in instances of low data availability to be showcased.

Following the creation of a suitable model, efforts must be made to design an effective operating strategy. To do so, sensitivity analysis is used to further elucidate process behaviour (Chapter 5). Herein, a range of glucose feeding schedules were tested *in-silico* to predict which process outputs were most sensitive to changes in the volume of glucose fed on each day throughout the cell culture. After, a dynamic optimisation study is presented using different scenarios to justify an appropriate process control strategy. To do so, objective functions were tested focusing on the minimisation of glucose fed, lactate, or osmolality, and the maximisation of viable cell density, or product titre. Finally, an advanced process control algorithm is deployed in a bioreactor utilising the previously developed mechanistic model and most appropriate operating strategy. More specifically, the feed of glucose to the cell culture is defined in real-time using non-linear model predictive control with an economic multivariable objective function. This allows for multi-output control simultaneously reducing the extracellular concentration of lactate, minimising the volume of glucose fed, and minimising osmolality, all whilst achieving the desired final product titre. This is then compared to open-loop optimisation with the same objective function and constraints that were used in the advanced process controller. All results presented in this thesis are summarised in the final chapter (Chapter 6) alongside the key limitations and future research required towards enabling process automation in upstream biopharmaceutical processing.

Chapter 2 Background & Literature Review

2.1 Overview of Monoclonal Antibodies

Monoclonal antibodies are the predominant class of biopharmaceuticals and have “revolutionised medicine” (Mullard, 2021; Vargason et al., 2021). In the immune system, antibodies are a protein secreted by B-lymphocytes as a response to identify and destroy foreign pathogens. Because of their high affinity, antibody-based treatments present the opportunity for extremely targeted medicines. For example, in cancer treatment, monoclonal antibodies can be used to target a specific antigen that is over-expressed on cancerous cells and relatively under-expressed on healthy cells. In doing so, the cancer cell is targeted for destruction by phagocytic cells (Zahavi & Weiner, 2020).

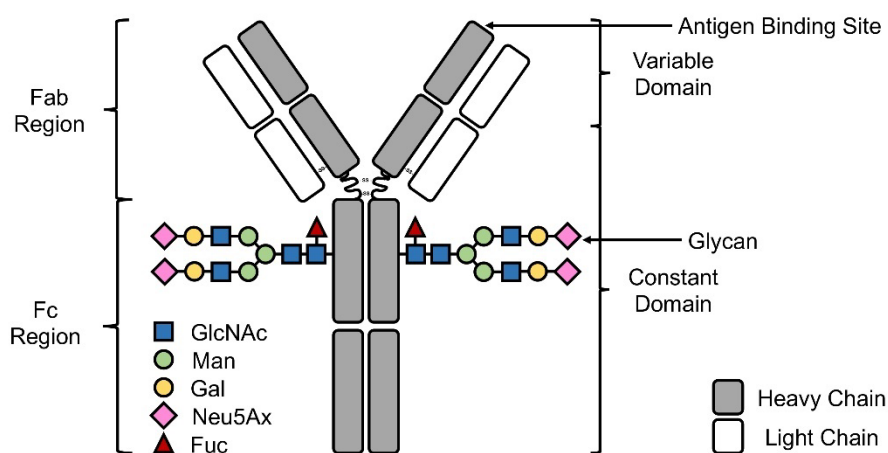


Figure 1 Structure of an IgG antibody.

Structurally, antibodies consist of pairs of identical heavy and light polypeptide chains and can be classed into different isotypes depending on the type of heavy chain. The most common class of antibodies in biopharmaceutical products are immunoglobulins G (IgG) which have a well-known Y-structure as shown in Figure 1. Other classes are IgA, IgD, IgE, and IgM

(Sommerfeld & Strube, 2005). The antigen-binding (Fab) region of the antibody is responsible for the recognition of antigens. The crystallisable (Fc) region is responsible for mediating interactions between the antibody and other parts of the immune system (Buss et al., 2012). Oligosaccharides, known as glycans, can attach to both the Fab region and the Fc region through glycosylation. However, N-linked glycosylation in the Fc region is the most common, and of importance in biopharmaceutical processing due to its significant influence on the effectiveness and stability of the final product (Liu, 2015).

The technology paving the way towards the production of monoclonal antibody treatments at scale was discovered in 1975 by César Milstein and Georges Köhler (Köhler, G., Milstein, 1975). Initially, this technology was met with doubt with a patent response stating that it was difficult “to identify any immediate practical applications which could be pursued as a commercial venture” (Tansey & Catterall, 1995). In contrast, by 1986, muromonab-CD3 became the first mAb to be approved by the FDA for its use in preventing kidney transplant rejection (Steinitz, 2014). As of 2023, there have been a total of 1437 antibody-based therapeutics subjected to clinical trials as part of steps towards commercial production and distribution (Kinch et al., 2023). That said, the transition to this has not been in the absence of further scientific advancements described below mainly tackling the issues surrounding immunogenicity.

Originally, initial monoclonal antibody treatments were murine such as muromonab-CD3. Here, these treatments were foreign to the human immune system. Therefore, these murine antibodies suffered from immunogenicity due to the high level of non-human epitopes and the resulting human anti-mouse antibody response (Jarboe et al., 2014). Following this, in 1994, the first chimeric (human–mouse) monoclonal antibody was approved for heart disease, abciximab (Tamhane & Gurm, 2008). In chimeric monoclonal antibodies, the murine constant domain was replaced by human constant regions to make the antibodies 70% human (Chames et al., 2009). Overall, the first developed chimeric monoclonal antibodies showed extended half-life in the human body and reduced immunogenicity versus murine therapies (Buss et al., 2012). However, the risk of a patient producing anti-drug antibodies when using chimeric-based treatment was still too high (Buss et al., 2012). To reduce this risk, humanised monoclonal antibodies were developed by substituting the heavy-chain variable region of a chimeric antibody onto a human antibody framework. In doing so, these monoclonal antibodies were 85-90% human with improved immunogenicity (Chames et al., 2009). The first

humanised monoclonal antibody (daclizumab) was approved for transplant rejection by the FDA in 1997. Further advances in antibody engineering eventually enabled the development of fully human monoclonal antibodies such as adalimumab (approved by the FDA in 2002) and panitumumab (approved by the FDA in 2006). Fully human antibodies significantly reduce the risk of the patient developing anti-drug antibodies, although they may still develop immune responses.

Despite many advances, monoclonal antibodies as relatively large molecules with complex post-translational modifications such as glycosylation, require sophisticated expression systems. Further, since large dose amounts are often required to achieve clinical efficacy, the production of monoclonal antibodies at scale involves the use of large cell cultures followed by extensive purification under good manufacturing practice (GMP) and subject to strict regulations (Glassy & Gupta, 2014). This very briefly explains the high monoclonal antibody production costs, thus treatment costs limiting the accessibility of such therapies (Chames et al., 2009; Glassy & Gupta, 2014). There is ongoing research on the use of other expression systems. One example is using microorganisms such as *Escherichia coli* to produce monoclonal antibodies, but these systems have been mainly used to produce non-glycosylated antibody fragments (Spadiut et al., 2014). Due to the requirement for a product with minimal immunogenic side effects and for human-like post-translational modifications such as glycosylation, mammalian host cell lines, specifically, CHO cells, remain the predominant expression system to produce monoclonal antibodies.

2.2 Production of Monoclonal Antibodies

The production of monoclonal antibodies can be split into (a) upstream processing and (b) downstream processing as shown in Figure 2. During upstream processing, host cells are cultured through progressively larger bioreactors known as seed fermenters. When a suitable inoculum volume is available, cells are inoculated within a production fermenter where the bulk of the biopharmaceutical product is expressed. Following this, it is necessary to purify the product using a series of downstream processing unit operations. Here, many other impurities are removed such as nucleic acids, host cell proteins, and cell debris to ensure product safety (Papathanasiou & Kontoravdi, 2020). During upstream processing, product titres can now exceed 10 g/L (Y. M. Huang et al., 2010). Meanwhile, monoclonal antibody-based products have dosage concentrations of around 200 mg/L hence the main task in downstream processing is impurity removal (Das, 2016; Hu et al., 2020).

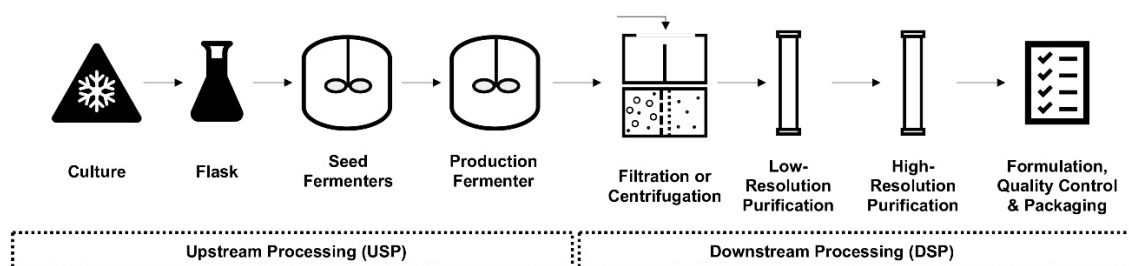


Figure 2 Schematic of a biopharmaceutical process.

The primary purpose of upstream processing is to produce a safe and efficacious product with consistency. Secondly, the product must be generated at a sufficient rate such that the overall process is economically viable (Walther et al., 2019). Previously batch mode was the main mode of operation for products licensed prior to the 1990s achieving final titres of around 100 mg/L (Jayapal et al., 2007; Ozturk & Hu, 2005). During batch processing, the limited amount of nutrients led to premature cell death caused by nutrient starvation. Now, fed-batch remains the most widely implemented mode of production achieving titres of around 5-10g/L (Walther et al., 2019). Here, a feed is continuously added to the cell culture at regular intervals, prolonging cell culture longevity and/or increasing product titre versus batch cell culturing. Finally, there is a great research interest towards perfusion wherein a product stream is removed

at the same rate as a feed stream is introduced. This could allow for highly prolonged cell cultures, lower capital costs, and the potential for reduced downtime (Sarna et al., 2023). However, perfusion has not been widely applied due to the complex process control required, large media requirements, and requirements for a genetically stable cell line (Walther et al., 2019). Meanwhile, despite advances towards continuous downstream processing (Papathanasiou et al., 2017), scheduling challenges could be created if only upstream is continuous whilst downstream is not.

2.3 Biopharmaceutical Process Development

Alongside product development, the process used to manufacture the final biopharmaceutical must also be capable of producing a product that ensures patient safety and product efficacy. As such various process development activities are performed at different scales as shown in Figure 3. In general, there is a strong interplay between selecting a suitable cell line for product expression, selecting suitable conditions tailored to that cell line, and producing a product with desired properties. In general, process development workflows are highly laborious and continuously iterated (for example, upon selection of a new product or cell line) and can increase time to market.

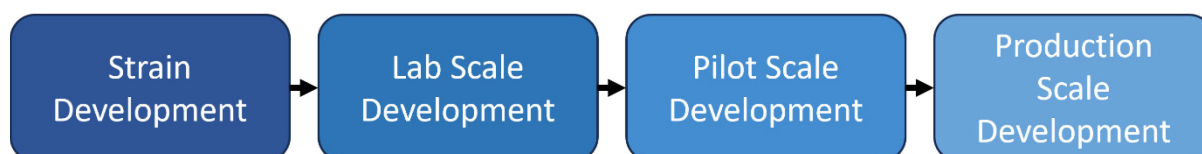


Figure 3 Process development workflow.

Initially, a high-performing strain is developed with a focus on cell line stability, high productivity, and favourable growth characteristics (Majumdar et al., 2024). Following this, fed-batch studies are performed to optimise medium, feed, and other process conditions for the cell line. In doing so, not only can the process performance be enhanced, but the necessary in-depth understanding of the cell line is built (J. Xu et al., 2022).

During process development, many different conditions must be efficiently screened. Traditionally, quality-by-testing (QbT) has been used whereby key product attributes are tested across the bioprocess and its development to ensure end-product quality (L. Zhang & Mao,

2017). However, QbT can lead to significant waste and may be inadequate for understanding the root causes of inefficiency across the process (Anurag S Rathore & Helen Winkle, 2009). This has prompted a shift over the last decade towards quality-by-design (QbD) which is now recommended by drug regulators such as the FDA and the EMA (I. Walsh et al., 2022). QbD involves ensuring a comprehensive understanding between critical quality attributes (CQA), critical process parameters (CPP), and critical material attributes (CMA) (Guerra et al., 2019; Maguire & Peng, 2015). These are defined as follows (Holm et al., 2017; Maguire & Peng, 2015; Rudge & Nims, 2017):

- CQA: "A physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality".
- CPP: "A process parameter whose variability has an impact on a CQA and therefore should be monitored or controlled to ensure the process produces the desired quality"
- CMA: " A physical, chemical, biological, or microbiological property or characteristic of an input material that should be within an appropriate limit, range, or distribution to ensure the desired quality of output material. "

QbD ensures that the targets of all CQAs are met. Further, QbD is already implemented within industry with the first full QbD approval in 2013 for Obinutuzumab (Sommeregger et al., 2017). Implementation requires the product target CQAs to be defined whilst the likely impact of CPPs and CMAs on CQAs is considered using a risk assessment (Looby et al., 2011). This is developed further using design of experiments, multivariate data analysis, or process modelling (I. Walsh et al., 2022). In doing so, a design space can be identified, and characterization studies used to determine acceptable variabilities of CPPs (Buss et al., 2012). Although not prioritised by QbD, some parameters can be classified as key process parameters (KPPs) which affect the process's key performance indicators (e.g., product titre) and must be controlled along with CPPs. Table 1 shows an example adapted from a recent study where QbD was used to define a process control strategy (J. Xu et al., 2022).

Table 1 Example process control strategy defined using QbD and adapted from (J. Xu et al., 2022)

Process Step	Process Parameters	Critical Process Parameters	Key Performance Indicators/Attributes
n-2 Seed	Temperature 35.5-37.5 °C Duration < 108 hours		Final viable cell density 4-10 x 10 ⁶ Viability ≥ 90%
n-1 Seed	Temperature 35.5-37.5 °C Duration < 108 hours		Final viable cell density 14-22 x 10 ⁶ Viability ≥ 90%
Production Bioreactor	Temperature 35.5-37.5 °C Initial Viable Cell Density 4-6.5 x10 ⁶ Duration 13-15 days	Cell generation <83	Final viability ≥ 50% Final Normalised Titre ≥ 1

It should be noted that CPPs can vary between processes as they are defined specifically for each using a risk assessment. In the example in Table 1, only the cell generation was considered a CPP. However, in other processes, variables such as pH and temperature could also be considered critical process parameters depending on the outcome of QbD analysis.

2.4 Process Measurement Technologies for Cell Cultures

To implement the control strategies defined during the final step of QbD, process measurement technologies are paired with suitable control algorithms (Anurag S Rathore & Helen Winkle, 2009). In this regard, the PAT framework, introduced by the FDA in 2004, has led to advancements in tools that can supply more timely process information aiming for increased process understanding and better quality control (FDA, 2004; Yan et al., 2024). However, there are still major challenges during upstream biopharmaceutical processes that make obtaining relevant process information difficult. Broadly, fed-batch cell cultures are rarely steady-state with time-varying dynamics that are very sensitive to change and thus require sensitive process measurement technologies (Teixeira et al., 2009; Vann & Sheppard, 2017). Several process

measurement technologies used during cell culturing are based offline and require sampling. However, sampling introduces a contamination risk to the cell culture. Therefore, the sampling frequency, thus measurement frequency, is constrained to reflect this e.g., to sampling and measurement once a day. On the contrary, measurement technologies based online must be able to withstand a range of harsh sterilisation procedures e.g., high temperatures from steam-in-place sterilisation. With these challenges, a sensor must provide high selectivity, sensitivity, robustness, repeatability, short response time, low detection limit and a long life (Biechele et al., 2015).

In general, three types of variables can be measured by sensors in bioreactors (Biechele et al., 2015):

- Physical variables such as temperature or pressure,
- Chemical variables such as pH, dissolved oxygen (DO) and substrates,
- Biological variables such as total and viable cell density/concentration, and extracellular concentrations of metabolites.

Sensors used to measure the variables can either record measurements (1) online, (2) in-line, (3) at-line, or (4) offline. During online measurements, sensors are placed in contact with the cell culture or within the bioreactor to offer real-time measurements. Alternatively, in-line measurements are taken by a sensor that is not placed directly within the reactor but outside where a sample flows to thus offering near real-time monitoring. Next, at-line measurements remove a sample volume from the bioreactor system for subsequent automated analysis. Finally, offline measurements represent the most manual solution where a sample is taken to be analysed manually away from the process. Many industrial fed-batch CHO cell culture processes rely on offline measurements where a major disadvantage is the delay between sample and analysis (Claßen et al., 2017). Overall, a summary of process measurement technologies used in fed-batch CHO cell cultures are shown in Table 2.

Table 2 Overview of key process parameters monitored by PAT during fed-batch CHO cell culturing, adapted from (Reyes et al., 2022).

Type	Monitored variable	Example process measurement technology	Controlled by	Online/in-line/at-line/offline
Physiochemical	Temperature	Thermocouple	Heating jacket	Online
	Dissolved oxygen	Optical	Sparging	Online
	pH	Optical	Acid/Base addition, sparging	Online
	Extracellular metabolite concentration	Enzymatic Assays	Media, feeding	Offline
Biological	Product titre	Turbidimetric Immunoassay	Media, feeding	Offline
	Cell concentration and characteristics	Capacitance, Microscopy	Volume adjustment	Online, Offline

In recent years, there has been a growing interest in applying spectroscopy for process monitoring of cell culturing with the aim of non-invasive continuous process monitoring. Specifically, there has been increased research interest in Raman and near-infrared (NIR) spectroscopy (Biechele et al., 2015). Initially, NIR lacked sensitivity and robustness with difficulties distinguishing low-concentration metabolites such as glutamine (M. Li et al., 2018; Sandor et al., 2013). However, improved use of regression models with NIR has shown promise for measuring non-glycosylated mAb concentration and glycosylated mAb concentration. of specific glycoforms (Zavala-Ortiz et al., 2020). That said, applications of infrared at a manufacturing scale are limited (Claßen et al., 2017). On the contrary, the number of applications of Raman is growing in both industry and academia (Claßen et al., 2017). Raman spectroscopy shows superior sensitivity and selectivity compared to other spectroscopy

techniques as it does not experience interference from aqueous solutions and can predict multiple metabolites simultaneously even at low concentrations (Abu-Absi et al., 2011; M. Li et al., 2018). Because of this, Raman has been applied across scales including 500L (Abu-Absi et al., 2011) up to 10,000L (Santos et al., 2019). However, a key disadvantage of Raman is the requirement of a model (e.g., partial least squares regression) to interpret the spectral data which can require recalibration upon changing process conditions (Mehdizadeh et al., 2015; Tanemura et al., 2023). Because of this and despite increasing interest, Raman is not yet widely deployed to fed-batch CHO cell cultures. Therefore, although advances in PAT are being made, a control strategy deployed at present would likely rely on existing widely available sensors including those based offline.

2.5 Influence of Conditions on CHO Cell Cultures

Major advances in tuning the conditions of CHO cell culture processes have been made to improve product titre, reduce by-products, and improve product quality (Coulet et al., 2022). The following section will give a brief overview of several key cell culture conditions and their effects on the process.

2.5.1 Impact of Temperature

Temperature is a key process parameter and potentially critical process parameter that significantly affects cell growth, productivity, and product quality (Z. Zhu et al., 2023). There are broadly two key strategies for temperature modulation within CHO cell cultures (1) keeping temperature constant e.g., at 37°C throughout the process and, (2) temperature shift during processing (McHugh et al., 2020). In the latter, reduced temperatures can be used to slow down cell metabolism. Before a temperature downshift, a cell culture is kept at its physiological temperature to maximise growth before the temperature is decreased to control cell proliferation and maintain a high viability over a prolonged period (Oguchi et al., 2006) Slowing down a cell's metabolism may prevent a build-up of inhibitory or toxic metabolites such as lactate and ammonia, postpone apoptosis, and improve product quality (Mellahi et al., 2019; A. Moore et al., 1997; Sou et al., 2015). Temperature downshifts have been used to increase specific productivity by up to 46% (W. C. Yang et al., 2016). In other studies, despite similar cell-specific productivities, a temperature downshift led to a 3-fold increase in final

product titre owing to prolonged viability (Mellahi et al., 2019). More recently, the effects of temperature downshift on metabolic pathways were explored with a balanced redox state being the overarching reason for improvements in cell culture performance (Le Gal et al., 2021; Z. Zhu et al., 2023). That said, the exact changes in performance due to a temperature shift are dependent on the cell line (McHugh et al., 2020). Further, the exact choice of temperature has shown to be influential with one study showing a downshift to 35°C decreased final titre significantly, whilst a downshift to 32°C increased final titres by 25% (McHugh et al., 2020). A suitable design of dynamic experiments (DoDE) may help uncover process-specific optimal conditions. However, in previous studies utilising DoDE, temperature setpoint was kept as a static input, rather than a dynamic input (Luo, Stanton, et al., 2023). Alternatively, a process model incorporating temperature, or temperature-dependant parameters, could be used to lessen the experimental burden and optimise culture conditions. (Sou et al., 2017).

2.5.2 Impact of pH

During CHO cell culturing, pH is typically maintained at around pH 7.0. As cells grow, glucose is metabolised to produce Adenosine-triphosphate (ATP) and lactic acid is produced as a by-product causing a decrease in pH. To balance this, a base is used (e.g., sodium hydroxide [NaOH] or sodium bicarbonate [NaHCO₃]) to increase the pH and return it to its setpoint. Further, an allowable range is generally set. The key advantage of doing this is to prevent excessive control action. In detail, although the base can neutralise the cell culture, it increases osmolality which can negatively impact cell growth (Alhuthali et al., 2021). As scale increases, mixing becomes more challenging due to the balance of maintaining a homogenous cell culture and preventing high shear stress from being exerted on the cell due to intense mixing. Similarly, pH can be increasingly difficult to control as base additions cause localised differences in pH for example, at the location of addition of up to +/- 0.8 in 8000L vessels (Langheinrich & Nienow, 1999). These challenges mean that the industry is shifting towards prioritising sparging for pH adjustment paired with a buffered media (Hoshan et al., 2019). The challenges of controlling pH when scaling up have led to several studies purposefully perturbing pH to observe the effects. For example, various setpoints for pH were tested on a CHO cell culture including low 6.7, 6.9 and 7.1 and determined that there was no difference between a pH of 6.9 and 7.1 with respect to cell growth (Lee et al., 2021). Similar indifferences at higher pH values were reported in a more recent study that purposefully introduced oscillations in pH between

7.1 and 7.5 and observed minimal impact on growth and productivity across different cell lines (Zakrzewski et al., 2022). At a lower pH of 6.7, a lower cell growth and thus a lower product titre was observed (Lee et al., 2021). Further, a lower pH also led to a threefold increase in the accumulation of toxic by-products such as ammonia (Lee et al., 2021). In addition, it was also observed that excessive additions of base led to increased lactate production (Jiang et al., 2018; Lee et al., 2021). That said, the effects of pH on cell cultures may be cell-line dependent. For example, an earlier study explored how variations in pH affected galactosylation whilst also discussing the mixed results present in other literature (Jiang et al., 2018). Therein, one cell line tested showed increased levels of galactosylation with increased deviations in pH but that was not reflected in the other two cell lines tested although the cause was not confirmed (Jiang et al., 2018). Ultimately, tight control of pH is crucial for improving process robustness as small deviations in pH may have negative consequences (Konakovsky et al., 2016; Michl et al., 2019).

2.5.3 Impact of Dissolved Oxygen

Dissolved oxygen (DO) is controlled to ensure oxygen availability within the cellular environment, enable cellular respiration and promote cell growth. Often this is set at 25-50% however some studies describe how fluctuations in DO have a greater negative effect on titre whilst the choice of setpoint is less important (Hippach et al., 2018). That said a setpoint too low (<10%) may lead to hypoxic conditions whereby cells may experience reduced growth rates (Kreuter, 2006). Further, limited DO could negatively affect galactosylation (Kunkel et al., 1998). On the contrary, setting a DO setpoint too high may lead to challenges in foam control and CO₂ stripping (Anane et al., 2021). At increased scales, heterogeneity in the cell culture can lead to low localised levels of DO that have been reported to prevent late-stage lactate consumption (Anane et al., 2021). However, therein DO did not affect productivity and, in general, research suggests that other process conditions, in particular feed, may be more influential of performance (Anane et al., 2021).

2.5.4 Impact of Feed

The variation in feed components to improve CHO cell culture performance has received significant interest (W. J. Xu et al., 2023). Feeds are generally complex with over 100

components typically classed into several categories including carbon sources, amino acids, inorganic salts, vitamins, lipids as well as many other additives (W. J. Xu et al., 2023). Unfortunately, most commercially available feeds are of proprietary composition. Further, there is no one-size-fits-all feed and unlike many other process industries where inputs to a system are dictated by closed-loop controllers, the exact feeding schedule, components, and volumes (or fed-batch recipe) of CHO cell cultures are often set pre-emptively and decided using empirical approaches (Mellahi et al., 2019). Each fed-batch recipe is process-specific, with effects dependent on the cell line, but broadly fed to reach or maintain a specific cell density, enabling the expression of a product of a desired specification. Finally, an effective feeding strategy should also consider downstream purification for example, by minimising the accumulation of by-products which later must be separated from the product.

Generally, all feeds are paired with glucose which is fed in bolus additions to the cell culture during fed-batch processing. Here, it acts as a key carbon source reportedly contributing 40-70% of the total carbon influx to the cells (Sellick et al., 2011). In detail, glucose is utilised in the glycolytic metabolic pathway to create pyruvate (Sellick et al., 2011). Pyruvate can then enter the tricarboxylic acid cycle to produce ATP, a key energy source required for cell function. However, during the cell's exponential growth stage, it is estimated that 75% of glucose consumption goes towards lactate production via pyruvate (Coulet et al., 2022). This represents a great limitation of CHO cells as lactate also inhibits growth (Lao & Toth, 1997; Martínez-Monge et al., 2019). As such, feed optimisation has been performed to identify alternative carbon sources in several studies e.g., using galactose (Altamirano et al., 2000). However, glucose is a key growth-limiting metabolite therefore, its substitution may result in a lower viable density (Altamirano et al., 2000). Other efforts to reduce metabolic by-products include the simultaneous feeding of glucose and lactate which was found to also reduce the extracellular concentration of ammonia, a toxic by-product, although growth and productivity remained similar (J. Li et al., 2012). More recently, flux balance analysis was used to explore feeding strategies that can induce efficient metabolism (Martínez-Monge et al., 2019).

Feed optimisation remains a large research topic with several hundred components constituting a single feed recipe. As such, the review of each component is outside the scope of this study with in-depth analysis available elsewhere (Coulet et al., 2022). More broadly, the effects of key process variables on product titre and quality are generally cell-line dependent and have also been recently summarised (Ha et al., 2022). Instead, the focus of this thesis will be on

tackling the *status quo* for reliance on large experimental approaches to optimise and control processes. Such approaches are costly, laborious, and add significant time to process development. Instead, there has been increased interest in the mathematical modelling of CHO cell cultures for *in silico* design and optimisation. With such models, advanced process controllers can be developed to directly control key performance indicators. Progress in each of these follows.

2.6 Modelling Mammalian Cell Cultures

2.6.1 Model Classification

Mathematical models are powerful tools that can significantly advance process development. Once developed, the use cases for process models include increasing process understanding, optimising feed and media (Maria, 2020), optimising process parameters (Sou et al., 2015), or process monitoring and control (Dewasme et al., 2023). Primarily, a model should provide an accurate description of cell culture conditions, whilst being able to extrapolate and make predictions of conditions not previously explored experimentally. Overall, an extensive range of model classifications have been proposed in literature (Figure 4) depending on the modelling approach, system characteristics, knowledge available, modelling domain, and modelling response (Tangirala, 2018; van den Bosch & van der Klauw, 2020). The following sections will review notable model classifications used within cell culture modelling applications.

Modelling Approach		System Characteristics		Knowledge Available		Modelling Domain		Model Response	
First Principles	Empirical	Linear	Non-Linear	Deterministic	Stochastic	Continuous	Discrete	Static	Dynamic
Parametric	Non-Parametric								
Single-Scale	Multi-Scale	Time-Varying	Time-Invariant	Black-Box	Grey Box	Time-Domain	Frequency-Domain	Lumped	Distributed

Figure 4 Mathematical model classifications. Adapted from (Tangirala, 2018).

The most widely compared model classifications when modelling cell cultures are mechanistic models, data-driven models, and hybrid models. Mechanistic models are built on first principles assumptions using extensive domain-specific knowledge (Kyriakopoulos et al., 2018) and have been reviewed extensively (Kyriakopoulos et al., 2018). For cell cultures, this can be in the form of differential equations to describe the mass balances of key cell culture components along with algebraic equations to describe metabolite utilisation and biomass growth rates (Mears et al., 2017). Mechanistic models require strong domain-specific knowledge and therefore necessitate significant time and resources to develop (Luo, Kurian, et al., 2021). Despite this, many mechanistic models often only approximate underlying assumptions about a process (van den Bosch & van der Klauw, 2020). This is especially the case for upstream biopharmaceutical processes (relative to downstream) whose mechanisms by large remain only partially understood due to the complexity of the cell metabolism and its intrinsic living nature (Narayanan, Sponchioni, et al., 2022). Even for known mechanisms, the translation of these into a mathematical domain remains challenging with some models built describing several thousand individual reactions (Strain et al., 2023). Despite the inclusion of such in-depth knowledge, these remain cell-line and thus process-specific (Strain et al., 2023). To date, the creation of generalized models for CHO cell cultures remains an open challenge and instead, models must be reevaluated when the cell line, process parameters or scale is changed. However, once developed, mechanistic models possess a key advantage owing to the injection of scientific assumptions which may lead to increased extrapolation capabilities (versus data-driven or hybrid) (Bayer et al., 2023). For the same reason, mechanistic models may also require less data to develop than other modelling approaches (Bayer et al., 2023).

Data-driven models are built empirically with less concern about whether the structure or parameters are biologically relevant provided the overall model allows for accurate prediction (Luo, Kurian, et al., 2021). The standardised structuring and parameterisation workflows give data-driven models an advantage over mechanistic models and can increase ease of development (Rashedi et al., 2023a). Further, data-driven models are also useful when model simplicity is preferred (Luo, Kurian, et al., 2021). For instance, if only analysis of the relationships between existing data is required, rather than a dynamic prediction of new process conditions, then statistical modelling is suited (I. Walsh et al., 2022; Yatipanthalawa & Gras, 2023). Previously, multivariate data analysis has been applied to elucidate the high correlation between lactate consumption during late-stage cell culturing and high productivity in fed-batch

CHO cell cultures (Le et al., 2012). Overall, a drawback of data-driven modelling techniques is that they can require greater amounts of experimental data versus mechanistic models (Narayanan, Sponchioni, et al., 2022). Further, data-driven models may have less extrapolation capability than mechanistic models (Herwig, 2018; von Stosch et al., 2014).

Balancing the benefits and drawbacks of data-driven and mechanistic models, hybrid models are increasingly being applied to CHO cell cultures partly due to the increased momentum of the industry to adopt principles of PAT (Doyle et al., 2023; von Stosch et al., 2014). Specifically, hybrid models can combine the interpretability of mechanistic model structures, whilst easing the development efforts of less known mechanisms that occur during cell culture (Narayanan et al., 2019). Hybrid models inherently allow for a large amount of structural flexibility (Narayanan et al., 2019). For example, mechanistic and data-driven aspects could be segmented in a serial approach (Thompson & Kramer, 1994) by modelling cell metabolism and growth mechanistically and feeding the results to an artificial neural network to model glycosylation (Kotidis & Kontoravdi, 2020). Alternatively, data-driven and mechanistic aspects can be more heavily integrated into a parallel approach (Thompson & Kramer, 1994) by using data-driven rate equations integrated into a series of ordinary differential equations (Doyle et al., 2023; Pinto et al., 2023). Despite the benefits, the main drawback is the increased amount of data that may be required for hybrid modelling relative to mechanistic models (Bayer et al., 2023). This may translate to increased experimental efforts and thus bring increased associated costs relative to mechanistic models. Ultimately, a balance must be drawn depending on data availability and the use case of the model.

Models can also be classed based on their modelling approach depending on the scales they incorporate (Figure 5). This could include modelling the cell culture at (a) the macro extra-cellular scale which may consider the bulk of the cell culture environment and assume homogeneity across the cell population (Sarna et al., 2023) ; (b) the meso-cellular scale for example whereby cells could be segmented into distinct populations with mass balances for each (Karra et al., 2010); or (c) the micro-intra-cellular scale which may consider reactions in a specific cellular compartment (Villiger, Scibona, et al., 2016).

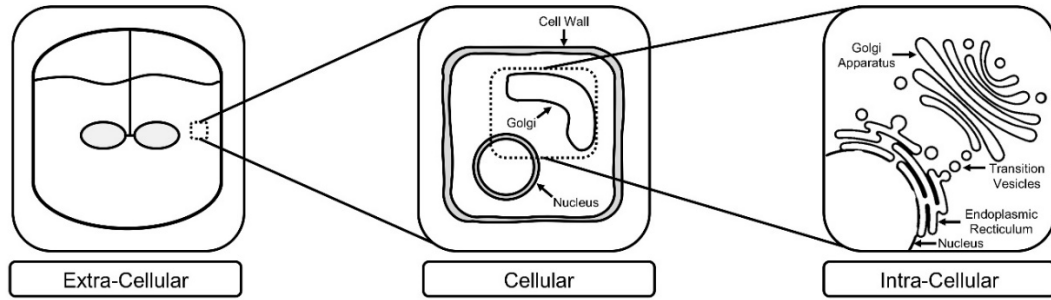


Figure 5 Scales across a cell culture.

In multi-scale models, several scales of the cell culture are modelled separately and brought together into one model for example to provide more accurate descriptions of cell metabolism (Karra et al., 2010). Multi-scale models could also model difficult-to-measure scales of the system (e.g. at an intra-cellular scale) using variables which are more easily measured (e.g., at an extra-cellular scale) and therefore there is potential for their use as a soft sensor to quantify hard to measure variables using more routinely available measurements (Villiger, Scibona, et al., 2016). However, increased modelling efforts may be required to describe a process across scales whilst the resulting model structure and parameters may have a large degree of uncertainty (Luo, Kurian, et al., 2021). Further, single-scale models that only approximate the behaviour of variables across one scale are a common dynamic model structure with suitable accuracy for many applications such as closed-loop process control (Rashedi et al., 2023a).

Not dissimilar from the classification of models by scale, models can be distinctly classed depending on the degree of segregation (Sidoli et al., 2004). In nature, cells are far from homogenous, hence segregated cell culture models aim to explore these inhomogeneities (Shirsat et al., 2015). Conversely, unsegregated models assume a homogenous cell population (Shirsat et al., 2015). Commonly across several types of mammalian cell cultures, segregated models are used to split the cell population based on cell cycle (Lam et al., 2008). However, a challenge with these models is the requirement of increased amounts of data to describe the subpopulations of cells. Intuitively, increasing advances in process analytical technologies will ease the development of segregated cell culture models. In the meantime, cell populations can, if desired, be modelled with sufficient accuracy by treating the population as one entity (Yatipanthalawa & Gras, 2023).

A further classification depends on the level of intracellular detail modelled. Structured models describe the inner structure of the cell such as with defined metabolic pathways, whilst unstructured models do not contain intracellular descriptions (Sidoli et al., 2004). For example, genome-scale models are a type of structured model built using knowledge of known intracellular reactions and used to design feeding strategies to improve product titre by up to two-fold (Fouladiha et al., 2020; Z. Huang et al., 2020). However, structured models require extensive knowledge of the cell mechanisms in some cases over 6000 reactions, as well as advanced analytical experiments to validate this knowledge (Strain et al., 2023). As such, reduced scale models have been developed to describe CHO cell metabolism with a few hundred important reactions and only a 2.5% reduction in prediction accuracy versus larger models (Jiménez del Val et al., 2023). Flux balance analysis is paired with genome-scale models commonly for post-hoc interrogation of experimental data. That said, there are two main limitations of current approaches, namely flux balance optimises an objective function that is commonly set to maximise biomass and thus can assume an overly optimal metabolism (Strain et al., 2023). Secondly, flux balance analysis alone outputs the steady-state of the cell culture which may be less useful for applications where dynamic capability is favourable e.g., for process control. Instead, dynamic flux balance analysis can be performed where the metabolite fluxes are estimated sequentially (Nolan & Lee, 2011). That said, dynamic flux balance analysis can be computationally intensive requiring optimisation at sufficiently short intervals to give accurate results. Alternatively, a structured kinetic model could be used for dynamic modelling (N. Chen et al., 2014). However, these models require significant parametrisation and structuring efforts.

Versus structured models, unstructured models describe extracellular variables. Previously, over 70% of unstructured mechanistic models for mammalian cell cultures were centred around a description of cell growth using Monod growth rate kinetics (Kyriakopoulos et al., 2018). Studies have shown the effectiveness of using Monod growth rate equations across scales (2L and 15L) and modes of operation (Craven et al., 2013). One of the key steps towards defining growth in mechanistic unstructured growth rates is the specification of growth-limiting (e.g., glucose, glutamine) and growth inhibitory (e.g., lactate, ammonia) metabolites (Karra et al., 2010; Xing et al., 2010). Previous authors have reviewed in depth the effect of each metabolite on growth (Pereira et al., 2018). However, growth can also be affected by other conditions and not solely metabolite concentrations. Therefore, several authors tackled this by including

modified pH-dependant Monod expressions, or temperature-dependant parameters (Gadgil, 2015; Sou et al., 2017). Crucially, careful choice of metabolites is required to ensure that the measurements of those selected are easily accessible especially if the model is to be used for closed-loop process control (Yatipanthala & Gras, 2023).

One of the great challenges with modelling CHO cell cultures (and more widely mammalian cell cultures) is that using approximations of a highly complex system, such as with unstructured unsegregated models, leads to a degree of uncertainty in the model structure and parameters (Kiparissides et al., 2011). Several developments have been made in research to deal with this. For example, sensitivity analysis can be performed on the parameters and model structure to quantify how parametric uncertainty affects the model outputs (Kiparissides et al., 2009). Further, insignificant parameters can be fixed whilst design of experiments can be utilised to further minimise uncertainty in significant parameters (Kontoravdi et al., 2010). As well as uncertainty stemming from the model, uncertainty in the system may stem from batch-to-batch variability in the inoculum, measurement errors, or variability in feeds. Many of the former modelling studies focus on deterministic modelling using a single value for parameters to deliver a single value output. Alternatively, stochastic models can handle uncertainties by providing a range of outputs and the relative likelihood of each. For example, Bayesian parameter estimation was used to model the seed train bioreactors using a mechanistic model wherein both an estimate of each unknown model parameter as well as the corresponding probability distribution were found (Hernández Rodríguez et al., 2019). In addition, a digital twin framework was used whereby as the process was operated, the model was updated to improve accuracy and reduce uncertainty (Hernández Rodríguez et al., 2019). Despite the promise of this method, stochastic models require increased computational effort to model and simulate versus deterministic models (Székely & Burrage, 2014).

Overall, over the past five years, there has been sustained interest in modelling CHO cell cultures. These advances are presented in Table 3.

2.6.2 Recent Advances in Modelling CHO Cell Cultures

Table 3 Recent advances in modelling CHO cell cultures. For research listed here, a literature review network has been built using [Inciteful](#) (follow link) which lists an up-to-date version of similar papers, review papers, predicts top authors, and predicts important papers (Weishuhn, 2024a).

Category	Manipulated Variable	Modelled Outputs	No. of Experiments (data split)	Culture Period (days)	Working Volume (L)	Reference
Hybrid and data-driven	Glucose feed and variations in the inoculum (Rouiller et al., 2012)	Xv, [mAb], [Glc], [Lac], [Gln], [Glu], [Amm], Osm, V	81 (8:2)	10	3.5	(Narayanan et al., 2019)
Mechanistic	Galactose feed (6.5mM-50mM) and uridine feed (0.76-20mM)	Xv, Xt, [Glc], [Asn], [Amm], [Lac], [Urd], [Gln], [Glu], [Gal], [Asp], [mAb], glycoforms, V	6	12	0.1	(Kotidis, Jedrzejewski, et al., 2019)
Hybrid	Various described in (Kotidis, Jedrzejewski, et al., 2019), (Villiger, Steinhoff, et al., 2016) and (Bydlinski et al., 2018)	Xv, Xt, [Glc], [Asn], [Amm], [Lac], [Urd], [Gln], [Glu], [Gal], [Asp], [mAb], Glycoforms, V	6 (5:1)	12	0.005-0.1	(Kotidis & Kontoravdi, 2020)

Mechanistic (Jimenez del Val et al., 2016) and data-driven	Temperature (33.5-37), pH (6.8-7), Feed	mAb, Xv, [Glc], glycoforms, V	5	14	0.05	(O. Yang & Ierapetritou, 2021)
Hybrid	Feed including glucose with disturbances from asparagine, glutamate, and copper feed.	Xt, Xv, [mAb], [Lac], [Amm], [Glc], [Gln], [Glu], [Ala], [Asn], [Asp], [Pro], glycoforms, V	15 (12: 3:1)	16	0.1	(Luo, Lovelett, et al., 2021)
Various mechanistic to data-driven	Feed of glucose and glutamine. Experimental data from various sources.	nAb, Xv, [Glc], [Lac], [Amm], [Gln], Osm, V	80 in-silico (20:20). 67 exp. (16:51)	14	3	(Narayanan, Luna, et al., 2022)
Mechanistic	Temperature (32-37°C), Feed	HCP, [mAb], Xv, Xt, cell volume, [Asn], [Amm], [Glc], [Lac], [Ala], [Gln], [Glu], Osm, [Na], [K]	4	14	1.2	(Alhuthali & Kontoravdi, 2022)
Hybrid and mechanistic	Feed time, feed rate, feed of glucose, feed of glutamine (Möller et al., 2019)	[mAb], Xv, Xt, [Glc], [Gln], [Amm], [Lac], [Gln]	33 (4:28, 19:14,33:0)	10	0.03-0.04	(Bayer et al., 2023)

Hybrid Mechanistic	Feed. See (Kyriakopoulos & Kontoravdi, 2014)	mAb, Xv, Xt, [Glc], [Lac], [Amm], [Ala], [Arg], [Asn], [Asp], [Gln], [Glu], [Gly], [His], [Ile], [Leu], [Lys], [Met], [Phe], [Pro], [Ser], [Thr], [Trp], [Tyr], [Val]	3	18	0.05	(Doyle et al., 2023)
-----------------------	---	--	---	----	------	-------------------------

In 2019, hybrid modelling was compared to data-driven modelling (Narayanan et al., 2019). Firstly, an artificial neural network with one layer was used to calculate specific rates and combined with a first principles model (Narayanan et al., 2019). Next, two partial least squares models were built with one model predicting the absolute values of the states, whilst the second model predicted the intermediate daily change in state (Narayanan et al., 2019). The hybrid model successfully predicted viable cell density based on error, whilst the statistical models both had an error that was greater than the experimental error (Narayanan et al., 2019). Interestingly, the ability of each model to extrapolate was tested by training the models only on the lower 50% of achieved endpoint product titres and using this to predict experiments for the upper 50% (Narayanan et al., 2019). On average, the hybrid model resulted in a 5-10% increase in error whereas the statistical models had up to a 45-50% increase in error (Narayanan et al., 2019). This model used data from previous research (Rouiller et al., 2012), however, it is rare to see and to have access to so many experiments as is visible (in Table 3) when compared to other studies. The data from these experiments was split 80:20 between training and testing data helping to prevent over-parameterisation (Narayanan et al., 2019). The large amount of data and data splitting could explain why the hybrid model performed exceptionally when extrapolating outside of the conditions upon which it was built.

In the same year, a mechanistic model was parameterised and subsequently used to maximise galactosylated antibody titre (Kotidis, Jedrzejewski, et al., 2019). Compared to the former study (Narayanan et al., 2019), the scale and number of experiments were reduced given the high associated costs and labour required (Kotidis, Jedrzejewski, et al., 2019). Using 7 experiments, a mechanistic model was built to describe the cell growth, metabolism, productivity, and glycoforms distribution (Kotidis, Jedrzejewski, et al., 2019). A Monod growth rate expression was used with glucose, asparagine selected as limiting metabolites, and ammonia, lactate, and uridine selected as inhibitory metabolites (Kotidis, Jedrzejewski, et al., 2019). Further, ammonia and uridine were used to describe cell death (Kotidis, Jedrzejewski, et al., 2019). It should be noted that some metabolites including uridine which was central to the growth prediction, required advanced analytical methods (chromatography) to determine their extracellular concentration (Kotidis, Jedrzejewski, et al., 2019). Further, nucleotide and nucleotide sugar quantification also required advanced analytics (Kotidis, Jedrzejewski, et al., 2019). This may cause challenges in applying the model for real-time control due to the difficulty, and length of time required to obtain state measurements. However, given that

galactosylation is a critical quality attribute that the model accurately predicts, this could be justified. On this part, excellent predictive capabilities were observed for viable cell density and product titre where for each the $R^2 > 0.95$ (Kotidis, Jedrzejewski, et al., 2019). Further, an accuracy of +/- 5% error was observed for most predictions of glycoforms distribution (Kotidis, Jedrzejewski, et al., 2019). Consequently, model-based optimisation was performed and a 90% increase in galactosylated antibody concentration was observed (Kotidis, Jedrzejewski, et al., 2019). Overall, this represents a significant contribution both in modelling, including of galactosylation, and an experimentally validated method using open-loop optimisation for process control.

The following year, an ANN was built to describe the glycan distribution with a mechanistic model built for the CHO cell metabolism (Kotidis & Kontoravdi, 2020). Using a manual search for hyperparameter tuning, 2 hidden layers were selected with 22 and 18 nodes respectively (Kotidis & Kontoravdi, 2020). Further, care was taken to avoid overparameterization by splitting the data for testing after which an informed decision was taken not to include a third hidden layer (Kotidis & Kontoravdi, 2020). The average absolute error of the glycoform distribution of the IgG was just 0.87% representing a substantial achievement (Kotidis & Kontoravdi, 2020). Further, robustness was tested by sequentially excluding inputs to the data-driven aspect of the hybrid model which resulted in a 1.25% error (Kotidis & Kontoravdi, 2020). Finally, a statistical model was built using PLS that reduced parameter and structuring efforts however the average error was 1.66% showing that the proposed hybrid model was superior to the proposed solely data-driven modelling (Kotidis & Kontoravdi, 2020).

Elsewhere, an unstructured kinetic model describing productivity and glycosylation that was previously developed (Jimenez del Val et al., 2016) was modified with additions on cell growth to capture temperature effects and on productivity to model pH effects (O. Yang & Ierapetritou, 2021). Following this, a data-driven model was developed to reduce the computational complexity of the model with the aim of easing process optimization efforts (Jimenez del Val et al., 2016). Overall, given the relatively low number of experiments, the data-driven model was able to predict the glycoforms with high accuracy and an error of <10% in all cases (Jimenez del Val et al., 2016). Next, the data-driven model was applied further to predict a design space that would ensure a glycoform would meet the requirements identical to that of Herceptin thus presenting a significant contribution towards demonstrating quality by design (Jimenez del Val et al., 2016).

Remaining with the trend of integrating glycosylation into models of CHO cell cultures, Luo et al demonstrated a workflow for adapting existing models to parameterised under one set of process conditions to a new set of process conditions (Luo, Lovelett, et al., 2021). Leave one out cross-validation was used with 12 experiments used to train a hybrid model and 3 experiments used for testing (Luo, Lovelett, et al., 2021). An interesting prospect would be to test whether this method of altering an existing model can be applied to other models. Further to what extent a process can be changed before this method becomes redundant e.g., due to plant mode mismatch. Overall, this presents a significant contribution as it relieves significant modelling efforts upon process changes.

Thus far, the key studies have generally compared a maximum of two model types. However, Narayanan et al proposed 7 models that lie between fully data-driven and fully mechanistic at varying “degrees of hybridisation” (Narayanan, Luna, et al., 2022). Therein the models were compared based on accuracy, model-building effort, extrapolation capability, and ease of utilisation (Narayanan, Luna, et al., 2022). Overall, models across the scale of hybridisation proposed have varying limitations. For example, data-driven models performed inaccurately and extrapolated poorly during instances of low data availability (Narayanan, Luna, et al., 2022). Interestingly, just 10 experiments were required for a hybrid model to perform similarly to a full data-driven model built with 50 experiments (Narayanan, Luna, et al., 2022). On the opposite end, a mechanistic model’s performance was observed to be independent of the number of experiments used to build it indicating that mechanistic models can be built with low experimental efforts (Narayanan, Luna, et al., 2022). However, mechanistic models introduce bias in the assumptions made during structuring and parametrisation and therefore the best-performing hybrid models outperformed all mechanistic models (Narayanan, Luna, et al., 2022). That said, only one fully mechanistic model structure was tested in this study thus perhaps the former could be an unfair conclusion against all mechanistic models. Other types of mechanistic models include structured vs non-structured, or segmented vs non-segmented as described in previous sections. This could also indicate why the mechanistic model did not offer the best extrapolation capability either. Although in this study mechanistic models were outperformed by some measures, there is an argument that mechanistic models may inherently drive advancements in process understanding. To counter this argument, this could also introduce bias and reduce the advantages promised by mechanistic modelling.

In a 2022 study, mechanistic modelling was used to further understand the mechanisms occurring during the secretion of host cell proteins during CHO cell culturing (Alhuthali & Kontoravdi, 2022). Therein, volume-based population balance modelling was used to link the negative effects of osmolality on growth rate, and the effects of hyperosmolality on cell volume (Alhuthali & Kontoravdi, 2022). The same model structure was used to produce two sets of parameters for differing temperature control strategies (Alhuthali & Kontoravdi, 2022). This model was particularly successful at modelling late-stage cell cultures and presents an opportunity for integration with downstream models towards end-to-end process optimisation (Alhuthali & Kontoravdi, 2022). Finally, the model was used for open-loop optimisation testing 6 objective functions for in-silico optimisation of product titre, to balance product titre with host cell protein secretion, and to predict the effects of a temperature shift (Alhuthali & Kontoravdi, 2022). The choice of objective functions is also discussed with respect to process design, especially considering downstream processing (Alhuthali & Kontoravdi, 2022).

In 2023, another comparative study was published between a pre-defined unsegmented mechanistic model and a hybrid model utilising an artificial neural network (Bayer et al., 2023). Therein, several different training and testing splits (see Table 3) were tested to explore the accuracy and extrapolation capabilities of each model (Bayer et al., 2023). The mechanistic model used Monod growth kinetics with glucose and glutamine as limiting metabolites, and ammonia as an inhibitory metabolite (Bayer et al., 2023). Meanwhile, the ANN was built using one hidden layer and the hybrid model was developed in the Novasign GmbH (Vienna, Austria) Hybrid Modelling Toolbox (<https://portal.novasign.at/>) (Bayer et al., 2023). Similar to previous comparative studies (Narayanan, Luna, et al., 2022), Bayer et al highlights how there is not a universal approach to the selection of an appropriate modelling workflow (Bayer et al., 2023). For low data scenarios, mechanistic models can make accurate predictions, but the benefit diminishes with increased data (Bayer et al., 2023). Further, the selection of appropriate experiments for model-building is important where previously, the optimal choice of model-building experiments led to over a 25% decrease in NRMSE versus the original training dataset (Bayer et al., 2023).

2.6.3 Challenges in Modelling Cell Cultures

In summary, knowledge gaps remain in the mechanisms driving CHO cell culture processes. As such, there has been a trend towards hybrid and data-driven modelling. However, the first major obstacle towards developing models of CHO cell cultures is the requirement for experimental data which comes with large costs and labour efforts. Given that multiple studies comparing hybrid to mechanistic models have highlighted the potential advantage of mechanistic models during instances of low data availability, mechanistic models still pose a clear advantage during such situations. In addition, mechanistic models must inherently be built in pursuit of increased process understanding. Crucially, during model-building, an awareness of the potential to introduce bias must be maintained especially during the structuring and parametrisation of unstructured, unsegmented mechanistic models. Other challenges in data collection pertain to the datasets often containing infrequent daily measurements and potentially misaligned time points. As such, a minor consideration during model-building should be towards more rigorous data analysis methodologies to tackle this e.g., automated data alignment, automated unit conversation, and automated regression as demonstrated in recent work (Doyle et al., 2023). Further, the latter is crucial before the development of any automated modelling workflows, including automated mechanistic modelling which remains an open challenge. Finally, demonstrations of workflows that limit the number of parameters and variables within a model to reduce over-parametrisation risk are of value (Doyle et al., 2023). On this point, it is also worthwhile to consider parametrisation strategies, recently reviewed elsewhere (Reddy et al., 2023). Comparison of parametrisation strategies is also a gap in many modelling studies.

2.7 Overview of Control of Mammalian Cell Cultures

Process control is the final step of implementing quality by design to biopharmaceutical processes (Anurag S Rathore & Helen Winkle, 2009). However, methods currently implemented for control of upstream biopharmaceutical processes significantly lag relative to other process industries (Mitra & Murthy, 2022). Broadly, upstream biopharmaceutical processes present challenges stemming from the production of a product from a heterogenous cell culture that has high sensitivity to bioreactor conditions (Luo, Kurian, et al., 2021).

Specifically, these processes often have highly non-linear, time-varying, and uncertain dynamics, paired with highly coupled process inputs (Dewasme et al., 2023). These challenges make it very difficult to implement robust process control methods capable of maintaining optimal process conditions (Nikita et al., 2023).

2.7.1 Control Problem Definition

Broadly, control systems can be operated in an open-loop or closed-loop manner (Mears et al., 2017). Controlled variables include (1) bioreactor conditions such as pH, DO, temperature, and extracellular glucose concentration, some of which may be critical process parameters, (2) key performance indicators such as product titre, and (3) critical quality attributes such as glycosylation. To control those variables, manipulated variables commonly used include base or acid additions, feed additions, or sparging. Overall, an effective control strategy will create conditions for high process performance and process robustness wherein a process will meet specifications even in the event of the numerous disturbances that are present in cell culturing.

2.7.2 Control Algorithm Classification

The most basic form of control is a system operated in an open-loop manner whereby predetermined process inputs are implemented as shown in Figure 6.

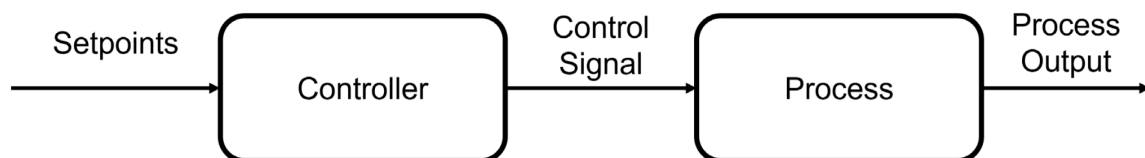


Figure 6 Open-loop system.

A key feature is that this control strategy does not rely on a feedback loop, and therefore does not necessitate advanced process sensors to achieve the desired process outputs. Because of this, this method is low cost, and simple to implement versus other control strategies discussed (Bolmanis et al., 2023). However, open-loop systems may not account for process deviations such as those created by process disturbances e.g., variations in the inoculum (Hernández Rodríguez et al., 2019). That said, open-loop systems are commonly implemented within the

biopharmaceutical industry using a fed-batch recipe with defined volumes of feed to be fed at fixed times (Bolmanis et al., 2023; Teixeira et al., 2009). Further, many industrial fed-batch systems are operated in some form of open-loop manner (Bolmanis et al., 2023). Examples of open-loop systems in literature are presented in Table 4.

Table 4 Examples of open-loop systems deployed for mammalian cell culturing. All defined using model-based optimisation.

Manipulated Variable	Control Variable	Controller Testing	Scale (L)	Expression System & Product	Reference
Temperature	Product titre	Experimental	0.05	CHO, interferon- γ	(Fox et al., 2004)
Feed glucose and glutamine	Product titre	Experimental	0.01	Hybridoma, mAb	(Lam et al., 2008)
Feed of glucose and glutamine	Various: (1) specific productivity, (2) product titre, (3) viable cell concentration, (4) integral viable cell concentration	In-silico	0.2	Hybridoma, mAb	(Koumpouras & Kontoravdi, 2012)
Feed glucose and glutamate	Product titre	Experimental	0.02	GS-NS0, mAb	(Kiparissides et al., 2015)
pH	Product titre	Experimental	1L	CHO, mAb	(Hogiri et al., 2018)

For open-loop systems manipulated variables can be defined by using a design of experiments. However, there are several examples where model-based strategies have been used to define manipulated variables for open-loop systems in literature. In 2004, research wherein an unstructured mechanistic model of a CHO cell culture was built with two sets of temperature-dependant parameters for 32°C and 37°C respectively and a structure centred around a Monod growth rate (Fox et al., 2004). The model was able to predict the productivity based on the day of the temperature shift and was subsequently validated in an open-loop manner leading to an improvement in productivity by 40% versus operating at the low temperature of 32°C, and 90% higher than operating at 37°C. This study represented one of the first model-based control strategies to manipulate the temperature shift of a CHO cell culture. Following this, in 2008, dynamic optimisation was performed on a previously developed mechanistic model (Kontoravdi et al., 2005) segregated based on cell cycle where model based open-loop optimisation led to increased cell culture lifetime and 40% increase in product titre (Lam et al., 2008). Further advances were made following this where in 2012, a model of a mammalian cell culture was utilised to test in-silico four different objective function for model-based dynamic optimisation as listed in Table 4 (Koumpouras & Kontoravdi, 2012). From these, it was postulated that as the predicted specific productivity was identical for all four objective functions tested, including where this was maximised, the limiting factor was the cell line (Koumpouras & Kontoravdi, 2012). Meanwhile, maximising viable cell concentration led to lower product titres than other objective functions (Koumpouras & Kontoravdi, 2012). Therefore, the authors concluded that integral viable cell density and product titre were the most appropriate variables to include within the objective functions for model-based optimisation (Koumpouras & Kontoravdi, 2012). This study is particularly interesting as it considered several objective functions rather than solely product titre (Koumpouras & Kontoravdi, 2012). In subsequent studies, Kiparissides et al. defined a mechanistic model of a GS-NS0 expression system that accounted for lag phase following inoculation which was then used to optimise the feeding frequency rather than using fixed feeding times. By varying the feed schedule of glucose and glutamate, a 30% increase in product titre was achieved (Kiparissides et al., 2015). Unlike the previously presented methods in this section, Hogiri et al., manipulated pH to control product titre (Hogiri et al., 2018). Further, this study was performed in 3L bioreactors (1L working volume) and therefore a closer representation of large-scale manufacturing conditions vs the flasks that were used for experiments in previous examples. In detail, a mechanistic model was built with 4 sets of parameters representing each

pH of 6.6, 6.8, 7.0, and 7.2 respectively (Hogiri et al., 2018). Following this optimisation was performed to maximise product titre and determine optimal pH shift. Experimental validation resulted in a 40% increase in product titre (Hogiri et al., 2018). The main challenge in the former study would be to achieve such tight control of pH in practice. Although perhaps dated, previous research suggests that at larger scale, pH can vary by 0.8 due to poor mixing (Langheinrich & Nienow, 1999). More recent research suggested that CHO cell cultures were not affected by small pH changes between 6.9 and 7.1, although this may be cell line dependant (Lee et al., 2021). Overall, open-loop systems have demonstrated success with strong experimental results when using model-based optimisation to maximise titre.

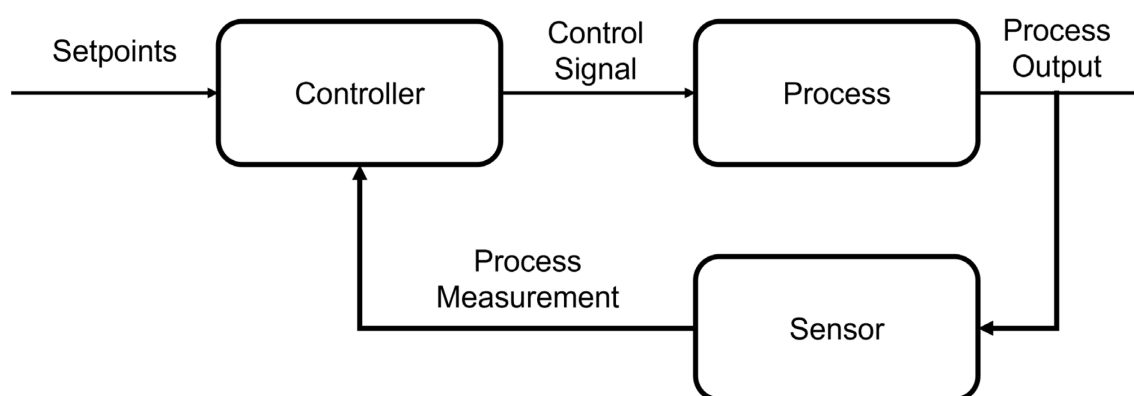


Figure 7 Closed-loop control.

Open-loop process control is well accepted in the biopharmaceutical industry because of its inherent simplicity and is broadly considered a “safe” operation strategy (Teixeira et al., 2009). For model-based control, if the process model is sufficiently accurate and robust, then open-loop systems may be sufficient to achieve reasonable control objectives as demonstrated by the success of studies in Table 4. However, in the presence of significant process deviations for example due to disturbances or plant model mismatch, open-loop systems may not perform adequately (Mears et al., 2017). Instead, closed-loop control (Figure 7) can help address these issues. Contrary to open-loop systems, closed-loop control systems rely on a feedback measurement prior to taking a control action thus there is a requirement for appropriate process analytical technology (Bolmanis et al., 2023). However, the advantage of closed-loop control is that it may lead to batch reproducibility by properly reacting to process deviations (Bolmanis et al., 2023).

One of the most common closed-loop control methods is proportional-integral-derivative (PID) controller shown in Figure 8. This is centred around correcting the process by reacting to the setpoint error, the difference between the setpoint and the measured output.

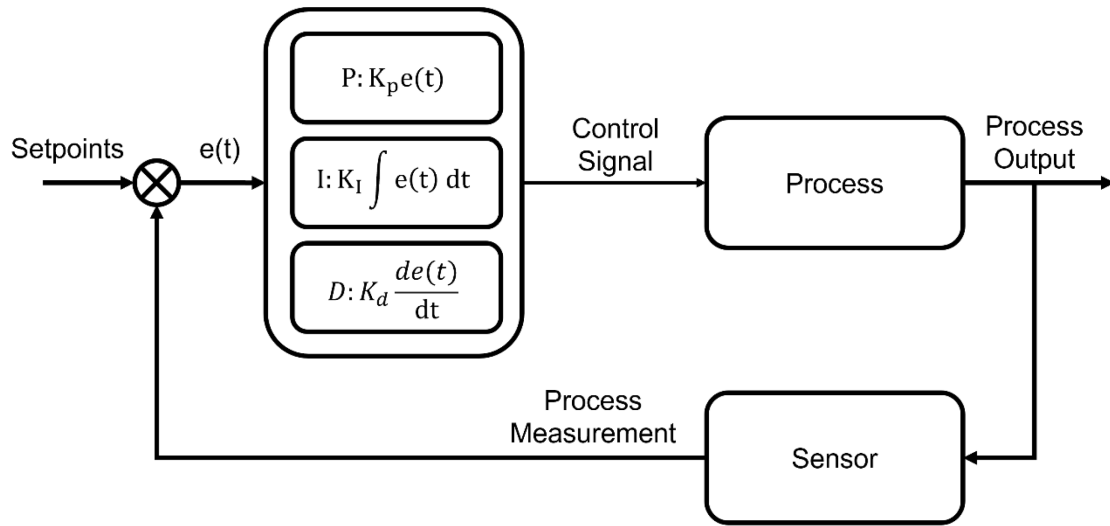


Figure 8 Process-integral-derivative control.

Using the setpoint error ($e(t)$), the control action is calculated based on Eq. 1.

$$u(t) = K_p e(t) + K_I \int e(t) dt + K_d \frac{de(t)}{dt} \quad \text{Eq. 1}$$

Here, $u(t)$ is the controller input at time t , K_p is the proportional gain, K_I is the integral gain, and K_d is the derivative gain. Briefly, the proportional term responds to the error at time, t , the integral term responds to the accumulated error aiming to reduce overshoot, and the derivative term responds to the rate of change of error aiming to reach setpoint faster. It should be noted that each controller gain can be tuned separately, and it is common to only use some terms e.g., P, PI, PD, or PID controller. For CHO cell cultures, PID controllers are generally utilised for regulatory control of base bioreactor conditions including pH, DO and temperature (Martens et al., 2014).

In 2011, Chotteau et al., demonstrated a PID tuning strategy to control DO in a 50L reactor (Chotteau & Hjalmarsson, 2011). Subsequently, DO and pH were simultaneously controlled in a 400L reactor (Chotteau & Hjalmarsson, 2011). In doing so, this research aimed to tackle one of the main challenges with scaling up where long mixing times create heterogeneities in the cell culture (Lee et al., 2021). When designing a controller at the 50L scale, random noise, and

variation in the oxygen consumption rate was added to test robustness of the controller under several scenarios. Meanwhile for the 400L scale, a water system was used to test several controllers and obtain satisfactory controller gains before being tested in a mammalian cell culture (Chotteau & Hjalmarsson, 2011). One downside of PID control is that it contains fixed parameters suited to linear control e.g., a P controller is suited to a process where the error is always proportional to the required corrective action. However, mammalian cell cultures have highly nonlinear process dynamics. Hence, adaptive controllers may be more suited and have been used for mammalian cell culturing. For example, Aehle et al., used an adapted proportional controller shown in Eq. 2 (Aehle et al., 2011).

$$u(t) = \left(1 + K_p e(t)\right) u_{SP}(t) \quad \text{Eq. 2}$$

Compared to a P controller which only takes the setpoint error ($e(t)$) as a variable, an adaptive controller can take several terms such as by using the desired set point trajectory, $u_{SP}(t)$. Aehle et al. used Eq. 2 to manipulate the feed of glutamine and control the cumulative oxygen uptake rate (tcOUR). The tcOUR was chosen as a controlled variable to suppress the effects of measurement noise that would otherwise be present in OUR. Several robustness tests were performed including deviating the initial viable cell density by 30% in which case the controller kept the process within $\pm 5\%$ relative deviation by 70h (Aehle et al., 2011) rather than by 110 h that was achieved without adaptive control (Aehle et al., 2012). Recent advances have led to possibility of using Raman for online measurement of the extracellular concentration of key metabolites such as glucose, lactate, or ammonia. As such, there is a renewed research interest in the use of PID for control of CHO cell cultures (Luo, Kurian, et al., 2021). One example of this is online glucose control using a proportional controller and Raman spectroscopy for feedback measurements (Matthews et al., 2016). Further examples are presented in Table 6.

Table 5 Examples of closed-loop control deployed for mammalian cell culturing.

Manipulated Variable	Control Variable	Controller Testing	Initial Working Volume (L)	Control Strategy	Reference
Sparging oxygen	Various: (1) DO, (2) DO, pH	Experimental	Various: (1) 50, (2) 400	PI control	(Chotteau & Hjalmarsson, 2011)
Feed glutamine	tcOUR	Experimental	2.5	Adaptive P control	(Aehle et al., 2011)
Feed glucose	Glucose	Experimental	5	P-Control	(Berry et al., 2016)

Model predictive control has been implemented widely across process industries and within other processes led to a range of benefits including a 5% throughput increase, a 15% reduction in operating costs, with 69% of users reporting process stability, and a reported 6–9-month payback time (Bauer & Craig, 2008a; William Y. Svrcek, Donald P. Mahoney, 2014). Yet, over 90% of biopharmaceutical processes do not use model predictive control and instead rely on classical control such as single input single output PID control (Mitra & Murthy, 2022). Despite this, model predictive control is seen as a promising method for closed-loop control of bioprocesses (Bolmanis et al., 2023),

In principle, model predictive control (Figure 9) requires a robust and accurate process model to anticipate the future behaviour of the system. Crucially, the success of the control strategy is heavily limited by the model robustness and accuracy (Schwenzer et al., 2021). For CHO cell cultures, modelling remains a challenge and can require large amounts of data (Bayer et al., 2023). This, and a lack of process validation of model predictive control applications in CHO cell culturing, partly explains why there is limited adoption (Mitra & Murthy, 2022). Once a model is developed, however, it is used to predict the process across a prediction horizon. Following this, optimisation is performed within a set of constraints to determine an optimal control action. Next, only the first control action is taken. Instead, upon receipt of new

data, the initial conditions of the model are updated, and optimisation is performed again to determine the next control action. This is repeated at every decision point to give a full control trajectory across the entire time horizon.

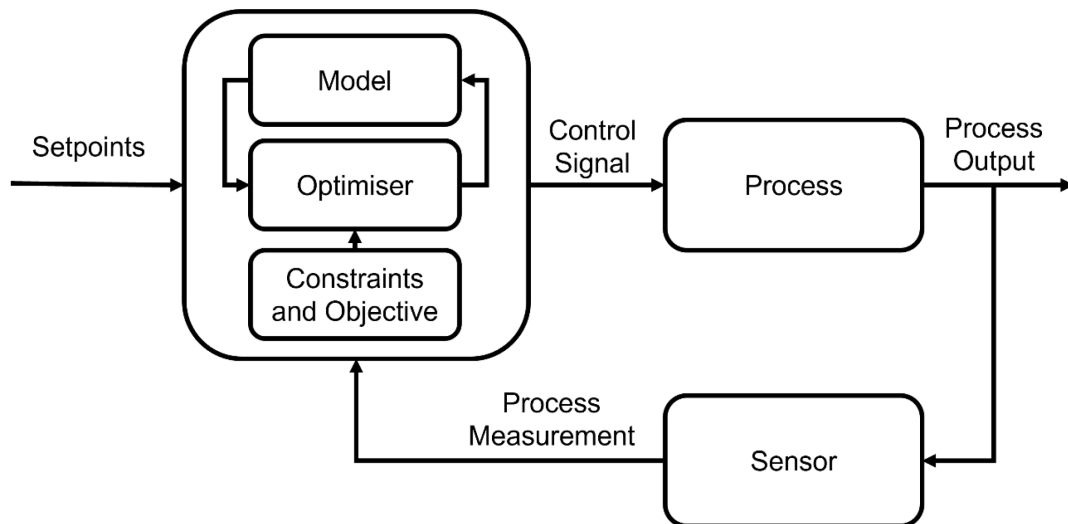


Figure 9 Model predictive control

A key advantage of model predictive control is the ease with which constraints can be set explicitly. Further depending on the model, model predictive control allows for implementation of multi-input multi-output control using inputs and an objective function. Common objectives include minimisation of error against a predefined trajectory. Alternatively, the minimisation or maximisation of process variables can also be set as an objective. Overall, due to the repeated re-prediction of the process, resulting in step ahead predictions, model predictive control may also offer robust performance through reduction in process-plant model mismatch. Examples of model predictive control are shown in Table 6.

Table 6 Examples of model predictive control deployed for mammalian cell culturing.

Manipulated Variable	Control Variable	Controller Testing	Initial Working Volume (L)	Reference
Feed of glucose and glutamine	Product titre	Experimental	2	(Teixeira et al., 2007)
Feed of glutamine	Total oxygen update rate	Experimental	2	(Aehle et al., 2012)
Feed of glucose	Extracellular glucose concentration	Experimental	15	(Craven et al., 2014)
Feed of mannose	High mannose	Experimental	0.002	(Zupke et al., 2015)
Feed of galactose	Galactosylated product titre	In-Silico	2	(Downey et al., 2017)
Feed of glucose, glutamate, arginine, aspartate, and asparagine	IVCD,	In-Silico	0.2	(Papathanasiou et al., 2017)

From Table 6, feed to the cell culture is often used as the manipulated variable for the control of product titre. Commonly this includes a major carbon source such as glucose or glutamine. For example, Teixeira et al. built a hybrid model of a baby hamster kidney (BHK) cell culture, whereby model parameters were updated in real-time (Teixeira et al., 2007). From this, a model predictive controller was used to determine the optimal feed of glucose and glutamine that would maximise product formation. During in-silico studies, this led to a predicted 10% increase in product titre and highlighted the importance of low initial extracellular concentrations of glucose and glutamine whilst viable cell density is low. Using the feed of

glutamine as a manipulated variable was subsequently tested experimentally by Aehle et al. on a CHO cell culture (Aehle et al., 2012). Here however, the total oxygen consumed was controlled, indirectly controlling viable cell concentration, with the aim of achieving process robustness rather than maximising process performance. To do so, model predictive control was used to minimise the error between the predicted trajectory of the oxygen consumed to a desired reference trajectory obtained during a previous experimental run. Overall, model predictive control led to a high degree of batch-to-batch reproducibility, although the exact RMSE deviations and other quantitative controller performance was not shown in this study.

Fixed setpoint tracking (rather than control variable minimisation/maximisation) was also used by Craven et al. whereby glucose feed was manipulated, and glucose concentration controlled in real-time. This was achieved by minimising the errors between the predicted extracellular glucose concentration and setpoint extracellular glucose concentration of 11mM at 0-100 hrs and 15mM after 100 hours (Craven et al., 2014). This study represents a major contribution as it uses a relatively small sample time of just 6 minutes, achieved experimentally by incorporating Raman spectroscopy for near continuous monitoring. As such, glucose can be supplemented at similarly short intervals potentially reducing osmotic stress that the cell culture is exposed to due to lower glucose feed volumes. Despite the use of Raman spectroscopy leading to increased measurement noise, the model predictive control strategy was able to effectively achieve near continuous setpoint tracking. Overall, this study demonstrates effective regulatory control of bioreactor conditions (e.g., glucose concentration). That said, closed-loop control of key performance indicators (e.g., product titre) or critical quality attributes (e.g., glycosylation) would be of value. Zupke et al. is one of the first to implement model predictive control experimentally to control glycosylation in a pilot scale (500L) reactor. Specifically, high-mannose N-linked glycosylation (“High Mannose”, %) is controlled by manipulating mannose feed and resulted in setpoint tracking achieving 6% High Mannose (+/- 1%).

As mentioned, one of the benefits of model predictive control is its ability for multiple input, multiple output control, thus is suited to CHO cell culturing. Papathanasiou et al. presents a notable in-silico example of multi-input control using the feed of several different metabolites to control antibody titre rather than solely glutamine and/or glucose. Further, varying sampling and feeding times during the fed batch schedule was also tested. Overall, model predictive

control may be suited to CHO cell culturing, and given the advances in process modelling described, the next step would be to implement these experimentally for real-time optimisation.

Over the past five years, there has been a growing interest in process control of CHO cell cultures with key studies shown in Table 5. Nearly 80% of studies listed use feed as a manipulated variable, with more 50% of studies specifically using glucose. As discussed in Section 2.5.4, glucose is the main carbon source used within the metabolism of CHO cells and therefore is highly influential on cell growth and product titre. In general, there remains significant opportunity for control of CQAs (e.g., HCP concentration) or KPIs (e.g., product titre) with over 50% of studies instead focusing on regulatory control for example of pH. Further, less than 10% of studies listed experimentally demonstrate closed-loop control of a KPIs or CQAs. Here, this research gap may be due to the challenges associated with development for use during controller design and deployment that describes these KPIs and CQAs. In addition, the challenges associated with experimental demonstration of process control of CHO cell cultures, including the associated high cost and labour efforts, may also explain why nearly 50% of studies listed focus on in-silico development of a controller rather than experimental demonstrations. In addition, for similar reasons this may explain why in all studies, controllers are designed on lab scale reactors (<15L). Overall, there is still a strong focus on fed-batch mode of operation which may be explained as this is the main mode of operation used at a manufacturing scale. Finally, despite the promised benefits of MPC described in previous sections, less than 50% of studies listed utilise MPC. Thus, there remains a significant research gap with only one study listed demonstrating MPC experimentally for control of KPIs or CQAs. A more in-depth discussion into relevant studies in control of CHO cell cultures is given in Section 5.1.

Table 7 Recent advances in control of CHO cell cultures. For research listed here, a literature review network has been built using [Inciteful](#) (follow link) which lists similar papers, review papers, predicts top authors, and predicts important papers (Weishuhn, 2024b).

Control Type	Manipulated Variables	Controlled Variables	Controller Testing	Mode of Operation	Scale	Reference
Open loop optimisation	Feed of galactose and feed of uridine	Galactosylated product titre	Experimental	Fed-batch	0.1	(Kotidis, Jedrzejewski, et al., 2019)
PID	Feed of glucose	Glucose concentration	Experimental	Fed-batch	2.5	(B. Moore et al., 2019)
MPC	Feed of a commercial feed	Lactate concentration	Experimental	Fed-batch	1-1.5	(Schmitt et al., 2019)
MPC	Feed of a commercial feed, pH	Product titre, viable cell density	In-silico	Fed-batch	Unspecified	(Hebing et al., 2020)
PID	Feed of lactic acid and glucose	Lactate concentration and glucose concentration	Experimental	Fed-batch	1.8-1.9	(Eyster et al., 2021)

PID	Feed of phenylalanine and glucose	Phenylalanine concentration and glucose concentration	Experimental	Fed-batch	5	(Webster et al., 2021)
PID	Bleed rate (flow rate out)	Viable cell density	Experimental	Perfusion	1.5	(G. Chen et al., 2021)
Reactive scheduling	Feed of galactose and feed of uridine	Galactosylated product titre	In-silico	Fed-batch	0.1	(Kotidis et al., 2021)
(1) PID and (2) MPC	Feed of a commercial feed including glucose	Product titre, glucose concentration	In-silico	Fed-batch	15	(Rashid et al., 2021)
Open loop optimisation	Feed of a commercial feed including asparagine, glucose and glutamate	(1) Product titre, (2) ratio of product titre to host cell protein concentration	In-silico	Fed-batch	1.2	(Alhuthali & Kontoravdi, 2022)
PID	Base flow rate, CO2 flow rate, stirrer speed, air flow rate, headspace, and O2 flow rate	Dissolved oxygen and pH	Experimental	Fed-batch	0.2-0.5	(Harcum et al., 2022)

PID	Feed of glucose and arginine	Glucose concentration and arginine concentration	Experimental	Fed-batch	2	(Domján et al., 2022)
PID	Feed of glucose	Glucose concentration	Experimental	Perfusion	1-5	(Graf et al., 2022)
Open loop optimisation	Feed of a commercial feed, sample rate, time of temperature switch, and time of operation mode switch to perfusion	Product titre, reactor run time, and glucose feed volume, working volume.	In-silico	Perfusion	8	(Kumar et al., 2022)
(1) PID and (2) MPC	Bleed rate (flow rate out)	Viable cell density	In-silico	Perfusion	10	(Pappenreiter et al., 2022)
PID	Feed of a commercial feed including glucose	Glucose concentration	Experimental	Fed-batch	2	(Webster et al., 2023)
PID	Bleed rate (flow rate out) and glucose feed rate	Viable cell density and glucose	In-silico	Perfusion	0.83	(Sakaki et al., 2023)
MPC	Feed of glucose	Glucose utilisation rate	In-silico	Fed-batch	0.05	(Monteiro et al., 2023)

PID	Feed of glucose	Glucose concentration	Experimental	Fed-batch	1	(Gibbons et al., 2023)
MPC	Temperature, pH, Feed of glucose, Bleed rate (flow rate out)	Product titre, harvest rate	In-silico	Perfusion	0.2	(Sarna et al., 2023)
MPC	Feed of a commercial feed including glucose	Viable cell density, total cell density, lactate, glucose concentration, glutamate concentration, lactate concentration, ammonia concentration, osmolality, glutamine concentration, sodium and potassium concentration	Experimental	Fed-batch	0.8-1.9	(Rashedi et al., 2023a)

(1) PID and (2) MPC	CO2 flow rate, and O2 flow rate	Dissolved oxygen, and pH	Experimental	Fed-batch	7	(Jones et al., 2023)
MPC	Feed of glucose and glutamine	Product titre	In-silico	Fed-batch	0.2	(Ohkubo et al., 2023)
MPC	Feed of a commercial feed including glucose, agitation setpoint, pH setpoint, temperature setpoint, DO setpoint	Viability, product titre, level of fucoslyation	In-silico	Fed-batch	0.3	(Luo, Kurian, et al., 2023)

Chapter 3 Project Aims

3.1 Research Proposal

In this thesis, several methodologies are developed computationally and deployed experimentally to advance towards process automation of upstream biopharmaceutical processing. Working with the predominant production process for the manufacture of monoclonal antibodies, this thesis uses a fed-batch operating mode along with an industrial cell line. Specifically, there is an aim to demonstrate effective closed-loop process control methods for fed-batch CHO cell cultures. To do so, a three-step approach is defined as follows:

1. Develop a process model,
2. Define an operating strategy,
3. Deploy a control strategy.

Overall, three process control strategies were tested experimentally including using manual bolus feeding of glucose, open-loop optimisation, and model-predictive control with the latter shown in Figure 10. Whilst previous studies have paved the way towards development of mechanistic models for fed-batch CHO cell cultures, in this study there is a purposeful development of mechanistic models for online application. As far as the author is aware, this presents the first attempt to deploy a mechanistic model within a model predictive controller for production of monoclonal antibodies to simultaneously control product titre, viable cell density, extracellular lactate concentration, osmolality and minimise feed.

3.2 Research Objectives & Proposal

To achieve the aim of this research, several research objectives were focused on using a mixture of computational and experimental methodologies shown in Table 8. These methods are demonstrated in two results chapters that are broken down as follows:

1. Chapter 4 Mechanistic Modelling of CHO Cell Cultures for Online Application:

The impacts of different glucose feeding regimes are investigated to develop a mathematical model of the fed-batch CHO cell culture with the following tasks.

- a. Three manual glucose feeding strategies are tested in a 2L single-use fed-batch CHO cell culture: a low glucose feed, a nominal glucose feed, and a high glucose feed.
- b. The cell culture is monitored throughout the cell culture collecting key data: (i) viable cell density, (ii) total cell density, (iii) product titre, (iv) key metabolite and by-product concentrations, (v) osmolality, (vi) temperature, (vii) pH, (viii) dissolved oxygen.
- c. An automated data analysis script is built in Python for unit conversions and key calculations such as growth rates and metabolite utilisation rates.
- d. Relationships between process variables are determined and a model structure and parameters that describe the fed-batch CHO cell culture is determined.

2. Chapter 5 Advanced Process Control of CHO Cell Cultures: Using the previous process model, in-silico experiments are used to define an operating strategy. Then, this operating strategy is deployed within a model predictive controller.

- a. Sensitivity analysis is used to describe the relationship between process inputs and process outputs.
- b. In-silico optimisation is used to rapidly test several process control strategies and reduce experimental burden.
- c. Model predictive control is deployed and compared to reactive scheduling during an operating error.
- d. Model predictive control is deployed and compared to open-loop optimisation and manual feeding during normal operation.

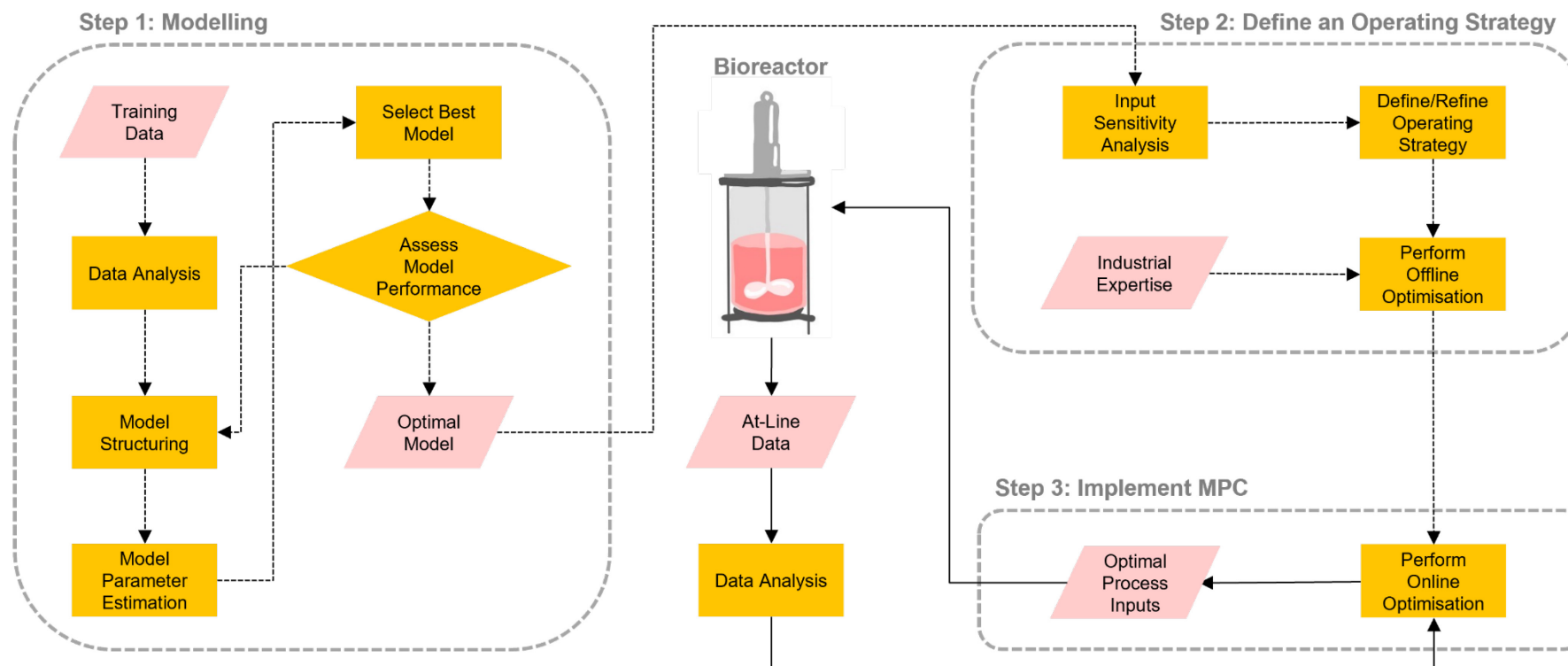


Figure 10 Towards implementation of model predictive control to fed-batch production of monoclonal antibodies.

Table 8 Summary of methods utilised within this thesis.

Thesis Chapter	Chapter Segment	Operation	Methods/Platform
Chapter 4 Mechanistic Modelling of CHO Cell Cultures for Online Application	Model training data collection using a manual glucose feeding	Cell culture	2Lsingle-use bioreactor (Sartorius Univessel® SU and Holder), monitoring and control hardware (Sartorius Biostat® B), monitoring software (Sartorius BioPAT® MFCS)
		Sampling/harvest	Centrifugation (Thermo Scientific™ Medifuge™ Small Benchtop Centrifuge)
		Product titre	High throughput analyser (Roche Cedex Bio HT® Analyser)
		Cell count	Cell counter (Beckman Coulter Vi-CELL XR Cell Viability Analyser)
		Metabolite analysis	High throughput analyser (Roche Cedex Bio HT® Analyser)
		Osmolality	Osmometer (Advanced Instruments OsmoPro Micro-Osmometer)

	pH	Blood gas analyser (Siemens Healthineers RAPIDPoint 500 Blood Gas Analyzer)
	Temperature control	Preset PID controller with fixed setpoint (35°C, Sartorius Biostat® B), heating jacket, temperature probe
	pH Control	Preset PID controller with fixed setpoint (pH 6.9-7.0, Sartorius Biostat® B), base (NaHCO ₃) addition (Sartorius Biostat® B), L-Sparger (Sartorius Univessel® SU)
	Dissolved oxygen control	Preset PID controller with fixed setpoint (50% DOT, Sartorius Biostat® B), electrochemical sensor (Sartorius Univessel® SU and Holder), L-Sparger (Sartorius Univessel® SU)
	Agitation control	Preset PID controller with fixed setpoint (225 rpm, Sartorius Biostat® B), 3-blade segment impeller 30° (Sartorius Univessel® SU), and motor (Sartorius)
	Feeding	Manual feeding at a fixed schedule
Modelling	Data analysis	Custom Python 3.9 script
	Model structuring	Manual

		Model parameterisation	Siemens gPROMS ModelBuilder 7.1.0
Chapter 5 Advanced Process Control of CHO Cell Cultures	Operating strategy development and deployment	Sensitivity analysis	Siemens gPROMS ModelBuilder 7.1.0
		In-silico optimisation	Siemens gPROMS ModelBuilder 7.1.0
		Real-time optimisation	Custom Python 3.9 script, Siemens gPROMS ModelBuilder 7.1.0
	Open-loop optimisation experiment	Same as model- building experiments.	Same as model-building experiments.
	Reactive Scheduling	Same as model- building experiments.	Same as model-building experiments.
	Non-linear model predictive control experimental	Same as model- building experiments.	Same as model-building experiments.

Chapter 4 Mechanistic Modelling of CHO Cell Cultures for Online Application

4.1 Introduction

Developing mathematical models of Chinese hamster ovary (CHO) cell cultures can aid process development towards defining a suitable operating strategy that would otherwise be determined through significant experimental efforts. In literature, there are notable examples of such models describing CHO cell metabolism including mechanistic (Kyriakopoulos et al., 2018), data-driven and hybrid models (Bayer et al., 2023; Doyle et al., 2023) as described in Chapter 2. Data-driven and hybrid models can require a larger quantity of data compared to mechanistic models (Bayer et al., 2023). However, during upstream process development, this is often not the case. The inclusion of first principles relationships in mechanistic models can make them suited to model processes with limited data (Bayer et al., 2023). In addition, even if hybrid/data-driven models can make more accurate predictions, the pursuit of biologically relevant mechanistic models inherently can advance process understanding by either validating existing mechanistic equations or postulating new mechanistic equations whose structure better reflects the underlying processes. The obvious major challenge here is the requirement for knowledge of the relevant mechanisms present. One prominent example in this respect is the complexity of modelling the extracellular concentration of lactate, especially where cells show a shift from production to consumption of lactate. Cell lines that display this mechanism are highly sought in an industrial setting due to its association with increased productivity (Rish et al., 2022). However, the metabolic shift from lactate production towards lactate consumption remains enigmatic (Hartley et al., 2018). This knowledge gap creates challenges during mechanistic modelling where lactate, as an inhibitory metabolite, is frequently placed at the

centre of models within Monod-based growth rate equations (Kyriakopoulos et al., 2018). Practically, lactate build-up could lead to decreased pH necessitating pH control for example with base addition thus increasing osmotic pressure on the cells. Interestingly, even hybrid models presented in recent research using relatively large datasets often describe extracellular lactate concentration less accurately relative to other metabolites (Bayer et al., 2023; Doyle et al., 2023). When combined with an incomplete awareness of mechanisms present during CHO cell cultures, there remains scope to demonstrate mechanistic model development with extremely limited datasets.

Another, open challenge for mechanistic models, is their purposeful development and subsequent application for real-time use. This presents unique challenges relative to existing models whose primary purpose is generally to increase process understanding prior to process operation. For real-time use, models should be developed to have high parsimony, minimising the number of variables to explain the process whilst ensuring goodness of fit. For example, it would be advantageous to ensure that models developed for real-time applications minimise their requirement for both quantity and complexity of process sensors or measurement technologies. In terms of complexity, process variables should be easily accessible, preferably measurable online or at-line. In terms of quantity, a model with many variables may rely on an increased number of sensors (hardware or software-based) to update the model and enable real-time integration. More so during process development, increased sensor numbers will cause increased burden in the event of sensor failures or inaccuracies (e.g., due to erroneous calibration) during which time the model may not be usable in real-time due to a lack of measurements. Greater sensor requirements also create the need for expanded creation and maintenance of data flows within a highly regulated environment. Yet, despite the above, the radical exclusion of process variables could subsequently hinder a model's ability to describe the highly complex and non-linear nature of CHO cell cultures.

Following this, Chapter 4 presents the development of a mechanistic model for use in real-time application, based on Monod-type equations to describe CHO cell growth, and material balances to describe extracellular concentration of key metabolites and product titre. Limited experimental efforts were used to generate training data.

Overall, Chapter 4 aims to investigate and provide insight into the following:

- Can a mechanistic model be used to describe CHO cell metabolism on an unseen cell line, with limited experimental data, suitable for real-time application?
- Can the logic of existing mechanistic modelling workflows be demonstrated and further formalised with the aim of being automated in the future?

4.2 Materials & Methods

4.2.1 Cell Culture

The experimental data used in this chapter for the construction, parameterization and testing of the model were generated through four experiments using a GSK mAb-producing CHO cell line. In preparation for the experiments, cells were revived from frozen cell banks into GSK proprietary growth media. Cells were incubated at 37°C, 5% CO₂, and 150 rpm in an orbital shaking incubator at a 25mm shaking diameter. Cells were sub-cultured every 3-4 days until sufficient volumes of cell culture were available for bioreactor inoculation at a suitable viable cell density. Each experiment contained cells for 5 passages to ensure the number of generations was kept close.

Four experiments were performed in 2L single-use bioreactors with a seeding density of $1 \times 10^6 \text{ cell.mL}^{-1}$ and a 1.6L working volume (Figure 11). Each process was operated under fed-batch conditions for 15 days with the feeding of a proprietary GSK feed on days 3, 6, 8, 10, and 13. Meanwhile, a supplementary feed of glucose (300 g.L⁻¹) was fed on days 3, 6, 7, 8, 9, 10, 13, and 14. During the model-building experiments, the volume of glucose fed was based on the cell population density and current extracellular glucose concentration measurements to deliver (1) manual feeding a nominal amount of glucose, MF-N, (2) manual feeding a low amount of glucose, MF-L, and (3) manual feeding a high amount of glucose, MF-H. The normalised volume of glucose fed in each experiment is shown in Figure 12. On day 10, to prevent glucose starvation due to a 3-day feeding gap at a high cell density, it is visible that the MF-L feeding strategy was fed more glucose than the MF-N. A repeat of the nominal amount of glucose (MF-N) was performed enabling calculation of standard deviation and an estimation of process variability with minimal experiments. PID controllers were used to maintain a fixed temperature of 35°C, and an allowable pH range between 6.9-7.0 controlled using NaHCO₃ and CO₂ additions. Dissolved oxygen tension (DOT) was controlled to a setpoint of 50% using air and O₂ supplied via a sparger at a variable ratio starting at 20 ccm of air with no pure O₂, after which, O₂ is controlled depending on demand and was capped at 300 ccm. Antifoam (Gibco™ FoamAway™ Irradiated AOF, ThermoFisher Scientific) was supplemented to the cell culture and a constant agitation speed of 225 rpm was applied throughout the process.

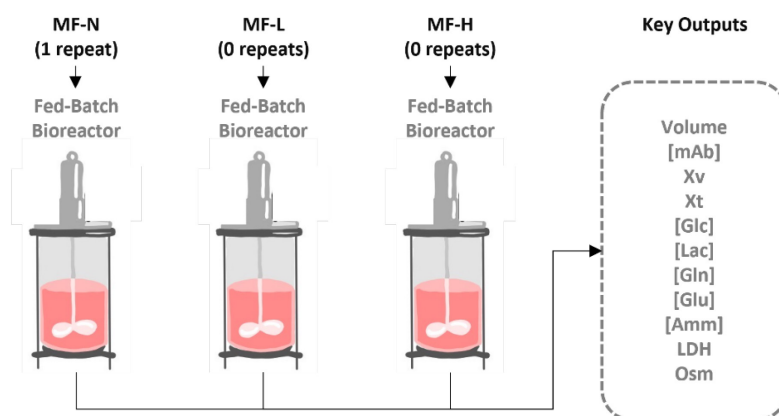


Figure 11 Model-building experiment overview with selected key outputs listed. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose.

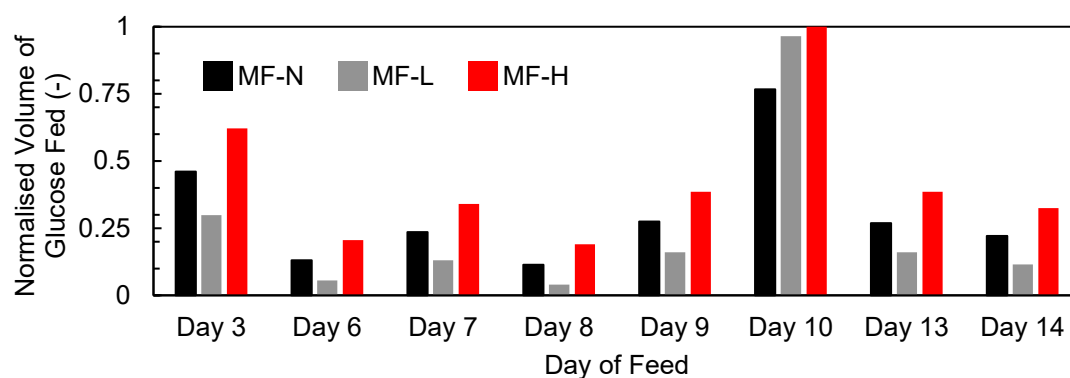


Figure 12 Normalised volume of glucose fed on each day throughout the model-building experiments. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose. Volumes are normalised based on the maximum volume of glucose fed which occurred in MF-H on Day 10.

4.2.2 Analytical Methods

Cell culture samples were taken daily. A RAPIDPoint 500 Blood Gas Analyzer (Siemens Healthineers) was used to calibrate and validate the pH sensor using a single daily measurement. To measure culture viability, viable cell concentration, and total cell concentration, 0.5mL of the cell culture sample was mixed with equal parts of TrypLE™

Express enzyme (Thermo Fisher Scientific) to enable the dissociation of aggregated cells. This was then split into three samples of 0.17mL and incubated for 10 minutes at 37°C before all three samples were measured by a Vi-CELL XR Cell Viability Analyser (Beckman Coulter) and the results averaged. The remainder of the cell culture sample was centrifuged for 2 minutes at 800 x g. The final supernatant was extracted, and the pellet was discarded. From this, a Cedex Bio HT® Analyser (Roche) was used with three repeat samples of 0.1mL of the supernatant to obtain extracellular concentrations of glucose, lactate, glutamate, glutamine, ammonia, lactate dehydrogenase and product titre. Finally, an OsmoPro Micro-Osmometer (Advanced Instruments) was used to obtain the osmolality measurement using two repeat measurements. The remainder of the sample was stored at -80°C.

4.2.3 Data Processing and Analysis

Following experiments, a Python script was built using the NumPy Python package to automate data processing using the workflow shown in Figure 13. Previously, this operation would be calculated manually for each data set which is challenging especially when feed schedules vary. Instead, this script allows for efficient analysis, especially for larger datasets with a high number of experiments. Similar approaches have been taken in previous research (Doyle et al., 2023) wherein a metabolite is assumed to be fed if the concentration increases more than 3 times during the process. However, in this research, the metabolites fed are explicitly defined by the user to enable ease of automation and prevent false assumptions, especially as this information is available prior to data processing.

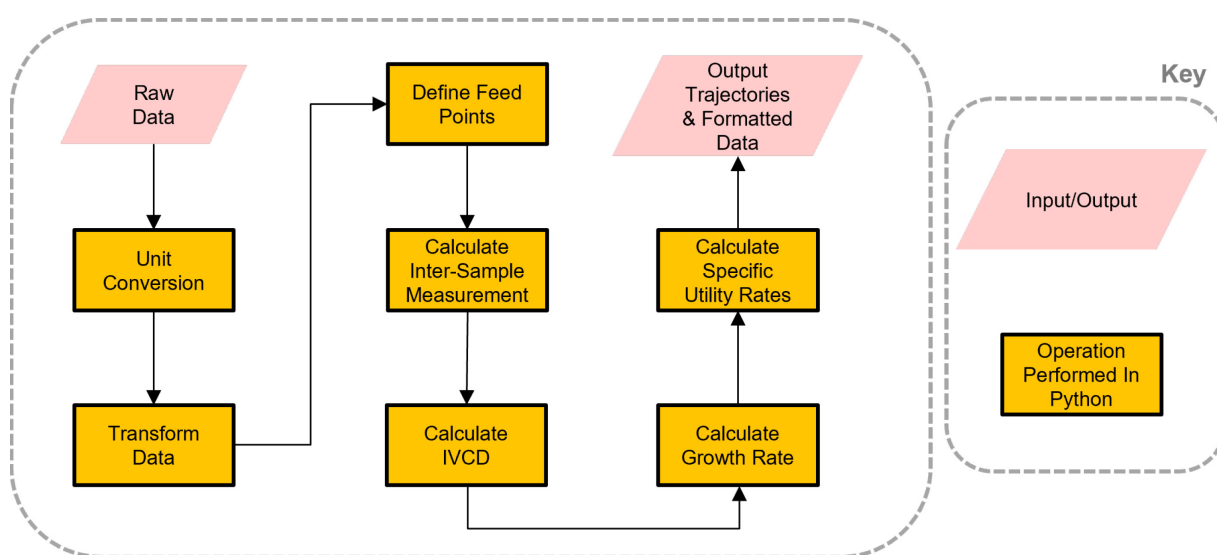


Figure 13 Automated data processing - workflow built in Python using NumPy.

From Figure 13, raw experimental data is inputted as a comma-separated value (CSV) file containing working volume, feed volumes, feed concentrations, extracellular metabolite concentrations, and product concentrations. Each variable must be labelled by the user prior to input according to the variable key shown in the appendix. In addition, feed points are labelled by the user prior to input to define when feeding occurs or when only measurements are taken. Following this, variables are labelled, and units are standardised across all variables as shown in Figure 14. This step ensures the code is generalisable which is crucial to enable efficient automation in further use cases. A dictionary is used containing 200 frequently used variable names, variable types, and allowable units. In addition, the molar mass of 14 key metabolites is stored and used for unit conversion between g.L^{-1} and mM. An example of how units are labelled and subsequently relabelled upon unit conversion from raw data to standardised data is shown in Table 9

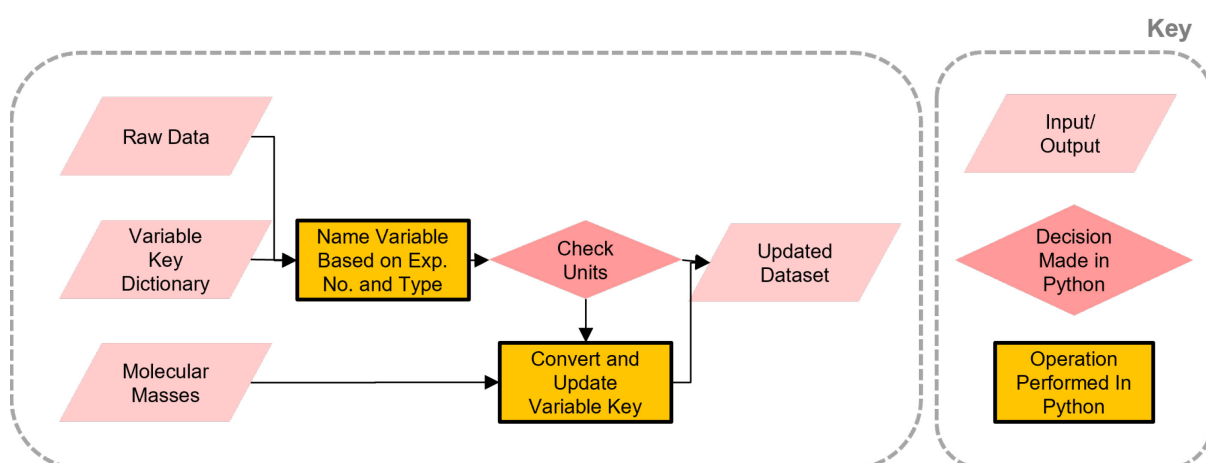


Figure 14 Automated data processing— unit conversion and variable labelling workflow performed in Python.

Table 9 Automated data processing— unit conversion example moving from raw data to standardised data.

	Raw Data	Variable Key	Standardised Data	Variable Key
Variable Name	4	Batch Time	4	Batch Time
Variable Type	2	Time	2	Time
Variable Units	2	Date Time	3	Hours
Data	01/01/2020 00:00		0	
	...		...	
	12/01/2020 00:00		288	

Data is structured as a long data structure where columns represent the variables, and batches are split by time points along rows. Further, feed times and metabolites are labelled to automate the step where the extracellular concentrations of metabolites in the cell culture are calculated after a feed is introduced using Eq. 3.

$$C_i(t) = \frac{(V_{in}C_{i,feed} + C_i(t-1)V(t-1))}{V(t)} \quad \text{Eq. 3}$$

Here, C_i (mM) is the extracellular concentration of metabolite and $C_{i,feed}$ (mM) is the concentration of metabolite in the feed. V_{in} (L) is the volume of feed fed in at the feed interval,

and V (L) is the working volume of the cell culture. Eq. 3 assumes insignificant metabolic effects during the feeding interval, starting at $(t - 1)$ and ending at (t) . Next, the integral viable cell density (IVCD) is estimated using Eq. 4.

$$IVCD(t) = IVCD(t - 1) + \frac{(X_v(t) + X_v(t - 1))\Delta t}{2} \quad \text{Eq. 4}$$

Eq. 4 uses the trapezium rule to estimate the integral where X_v (cells.L⁻¹) is the viable cell density. The estimated growth rate (μ_{growth} , h⁻¹) is calculated using Eq. 5 interpolated in-between the nearest measurements of X_v whilst the overall growth rate ($\mu_{growth,overall}$, h⁻¹) is calculated using Eq. 6. Finally, the cell-specific metabolite utilisation rates q_i (mM.cell⁻¹.h⁻¹) are calculated using Eq. 7.

$$\mu_{growth}((t - 1) + 0.5\Delta t) = \frac{\ln\left(\frac{IVCD(t)}{IVCD(t - 1)}\right)}{\Delta t} \quad \text{Eq. 5}$$

$$\mu_{growth,overall} = \frac{\ln\left(\frac{IVCD(t_{end})}{IVCD(t_0)}\right)}{\Delta t} \quad \text{Eq. 6}$$

$$q_i((t - 1) + 0.5\Delta t) = \frac{\Delta C_i}{\Delta IVCD} \quad \text{Eq. 7}$$

Despite what may seem like a rudimentary task, automation of data processing from experimental format to that suitable for *in silico* tools is sometimes overlooked and performed manually. Yet this forms the initial steps towards process automation and real-time application of existing *in silico* tools efficiently. The present work automates the calculation of metabolite utilisation rates from raw experimental data, further work could include automating constraint definition for flux balance analysis (FBA). This would allow for real-time FBA especially using recently published computationally efficient models such as CHOMPact (Jiménez del Val et al., 2023). Alternatively, further automation could build and include real-time process fingerprinting that was previously demonstrated as a tool for increasing process understanding (Sokolov et al., 2015).

4.2.4 Mathematical Modelling

A mechanistic modelling workflow shown in Figure 15 is utilised. Owing to their demonstrated performance during instances of limited data availability (Bayer et al., 2023). Data was processed using the function described in Section 4.2.3. Next, the growth rate equation and metabolite utilisation rates are defined based on experimental observations and data analysis described in Section 4.3 and Section 4.4.

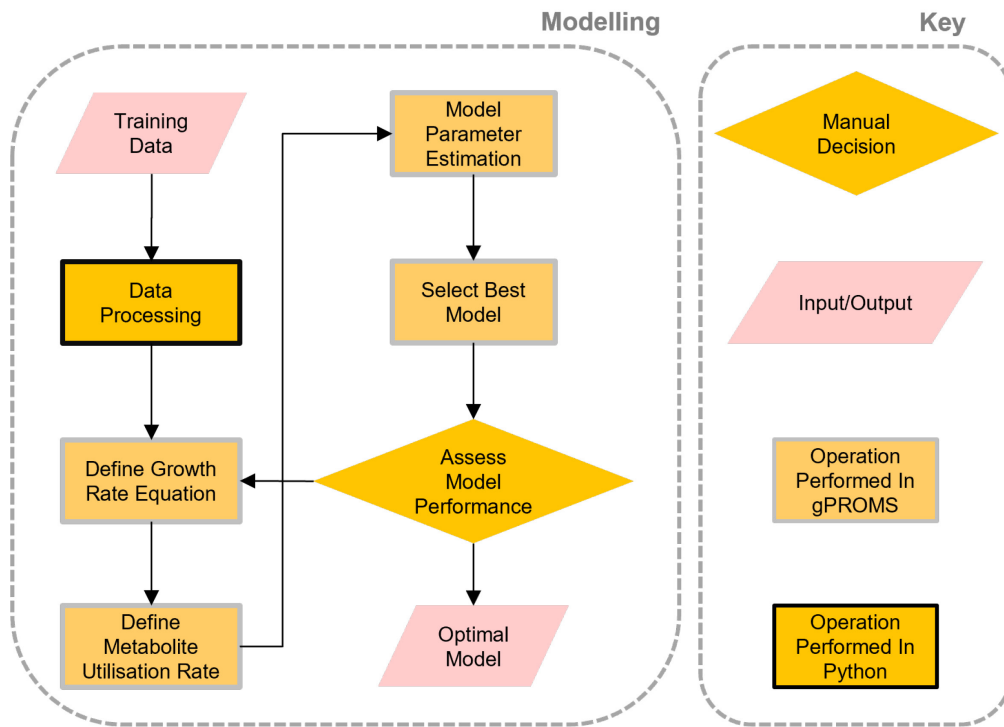


Figure 15 Mechanistic modelling workflow.

Model parameterisation was performed in gPROMS ModelBuilder v7.1.1 (Siemens PSE). Using Eq. 8, the maximum likelihood formulation was used for parameter estimation whereby a solver maximises the probability of predicting experimental data using parameters assuming normally distributed measurement errors and standard deviations.

$$\phi = \frac{N}{2} \ln(2\pi) + \frac{1}{2} \min_{\theta} \left(\sum_{i=1}^{NE} \sum_{j=1}^{NV_i} \sum_{k=1}^{NM_{ij}} \left(\ln(\sigma_{ijk}^2) + \frac{(\tilde{z}_{ijk} - z_{ijk})^2}{\sigma_{ijk}^2} \right) \right) \quad \text{Eq. 8}$$

Here, ϕ is the solver objective function minimised by optimising the physical parameters, θ subject to upper (θ_u) and lower (θ_l) bounds where $\theta_l \leq \theta \leq \theta_u$. N is the total number of

experimental data points, NE is the total number of experiments, NV_i is the total number of model variables measured at the i^{th} experiment, NM_{ij} is the total number of experimental measurements for the j^{th} variable in the i^{th} experiment, and σ_{ijk}^2 is the variance of the k^{th} measurement for the j^{th} variable in the i^{th} experiment. Finally, $\tilde{z}_{ijk}$ is the k^{th} experimental measurement for the j^{th} variable in the i^{th} experiment whilst z_{ijk} is the k^{th} model prediction for the j^{th} variable in the i^{th} experiment.

Once a dynamic model (described in Section 4.4) was parameterised it was tested by simulating the experimental conditions which were used to build it. Dynamic model simulation was performed in gPROMS ModelBuilder v7.1.1 (Siemens PSE) using the inbuilt DASOLV function to solve a set of mechanistic ordinary differential equations and algebraic equations. A discretisation interval equal to the smallest feeding interval was used and an absolute tolerance several orders of magnitude lower than the smallest parameter value or metabolite utilisation rate.

During model parameterisation, the uncertainty surrounding the values of the parameters may require large bounds. Further, the non-linearity of the system increases the risk that estimated parameter values may only be locally optimum. Sensitivity analysis can be used to decipher how uncertainty in a model can be apportioned to different model inputs. Uncertainty analysis is distinctly separate and only characterises uncertainty without identifying the accompanying responsible input factors (Saltelli et al., 2019). Using sensitivity analysis, input factors that contribute most to uncertainty can be explored further (factor prioritisation) to address and hopefully reduce it (Saltelli & Tarantola, 2002). For example, following factor prioritisation, design of experiments under parametric uncertainty can be used to further decrease model uncertainty (Nimmegeers et al., 2020). Sensitivity analysis can also be used to identify factors that can be fixed (factor fixing) if they do not contribute significantly to model uncertainty (Saltelli & Tarantola, 2002). This can simplify future parametrisation problems by reducing the number of parameters required to be estimated and thus the computational cost of parameterisation.

Global Systems Analysis within gPROMS ModelBuilder v7.1.1 (Siemens PSE) was used to explore the relative importance of each of the parameters used in the mechanistic model

(described in Section 4.4) and their influence on the mechanistic model outputs. This variance-based sensitivity analysis method is only suggested where the number of input factors, in this case, parameters, are under 20 (Saltelli et al., 2008), otherwise metamodeling may be required (Kucherenko & Zaccueus, 2021). Variance-based sensitivity analysis is used to determine first order and total effect indices for each parameter normalised by the variance in response (Saltelli et al., 2010). The first order effect for each parameter is attributed to the contribution to the uncertainty for that parameter alone and can be used for factor prioritisation (Saltelli & Tarantola, 2002), whereas the total effect is the combination of linear and nonlinear contributions from all parameters on a model's uncertainty and can be used to facilitate factor fixing (Saltelli & Tarantola, 2002). Quasi-random Sobol sampling is a sampling method that yields no overlapping points (Sobol, 1976) and was used to generate 4096 sets of parameters. Here, an increasing number of parameters were chosen until each analysis converged to the same values of sensitivity indices. These parameter sets were used during sensitivity analysis with bounds set dependent on the confidence intervals obtained during parameterisation, assuming a normal distribution for each parameter. The first order and total effect indices are estimates of the Sobol indices that converge to their true value as the number of samples increases (Saltelli et al., 2010). As such, the number of parameter sets must be sufficiently large such that the estimated Sobol indices remain stable.

4.2.5 Data Normalization

To protect GSK intellectual property, data was normalised prior to presentation using min-max normalisation described in Eq. 9.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad \text{Eq. 9}$$

4.3 Experimental Observations

Prior to modelling, experimental results should be analysed. Looking at Figure 16, despite significant differences visible in the extracellular glucose concentration of up to 50% between experiments, other measured variables do not appear to differ as substantially. This initial observation showcases an ideal characteristic of an industrial cell line which should be robust in the presence of process deviations.

Product titre trajectories can be observed in Figure 16 where the low glucose feeding strategy resulted in a 6.2% higher endpoint product titre than the nominal feeding strategy and a 25% higher endpoint product titre than the high glucose feeding strategy. This could be attributed to either (a) higher specific productivity, or (b) higher specific growth rate. The specific productivity shown in Figure 17 shows more variation than the specific growth rate shown in Figure 18. The spike in specific productivity in Figure 17 during late-stage culture of the high glucose experiment (MF-H) may be explained by a measurement error in product titre on Day 13 which appears off-trend relative to other measurements in Figure 16. Nonetheless, the peak viable cell density is also highest when less glucose is fed in MF-L and lowest in MF-H. Therefore, it can be observed that over-feeding glucose has a negative impact on growth which is also observed in previous research (Fan et al., 2015).

Interestingly, the death rate is also higher more frequently during low glucose feeding (MF-L) as can be seen in Figure 19 and is validated in Figure 16 where the dead cell density is highest in MF-L. In addition, increased death is backed up when the extracellular concentration of lactate dehydrogenase (LDH) is also highest in MF-L as can be seen in Figure 16. However, it can be observed in Figure 16 that both ammonia and lactate are lowest in MF-L and highest in MF-H. As such, the increased death rate in MF-L is not due to increased waste extracellular metabolites inducing increased apoptosis. Instead, during the stationary and death phase of cell culture, the extracellular concentration of glucose is approaching limiting levels and therefore this may cause nutrient starvation thus inducing cell death. In other research, the extracellular concentration of LDH is often used as an indicative measure of lysed cells (Chan et al., 2013; O'Mara et al., 2019) perhaps due to nutrient starvation.

Reduced availability of glucose may also induce lactate consumption (Martínez et al., 2013) which may also explain why the endpoint extracellular lactate concentration shown in Figure 16 is lowest in MF-L and highest in MF-H. The increased extracellular ammonia concentrations observable in Figure 16 in MF-H experiments could also indicate an inefficient metabolism by cells leading to over-production of ammonia. However, the higher extracellular glutamine concentrations of MF-H do not explain this. Similar to previous research where increased osmolality led to increased ammonia synthesis, perhaps other amino acids are playing a part in the slightly higher production of ammonia in high osmolality (Alhuthali et al., 2021). In this case, access was not available to all amino acid measurements, but measurements of these would allow a more comprehensive flux balance analysis to be performed to identify the effects on ammonia utilisation. The same research (Alhuthali et al., 2021) reported that high osmolality suppressed cell growth in CHO cells. The feeding regime in MF-L resulted in lower osmolality (Figure 16) compared to the other experiments, which may explain why MF-L is sustaining higher cell growth rates (Figure 18). Further, despite differences in osmolality, there is little observable difference in productivity between all three experiments as shown in Figure 17. This is validated in previous research which suggests hyperosmolality does not affect productivity in certain cell lines (Romanova et al., 2021).

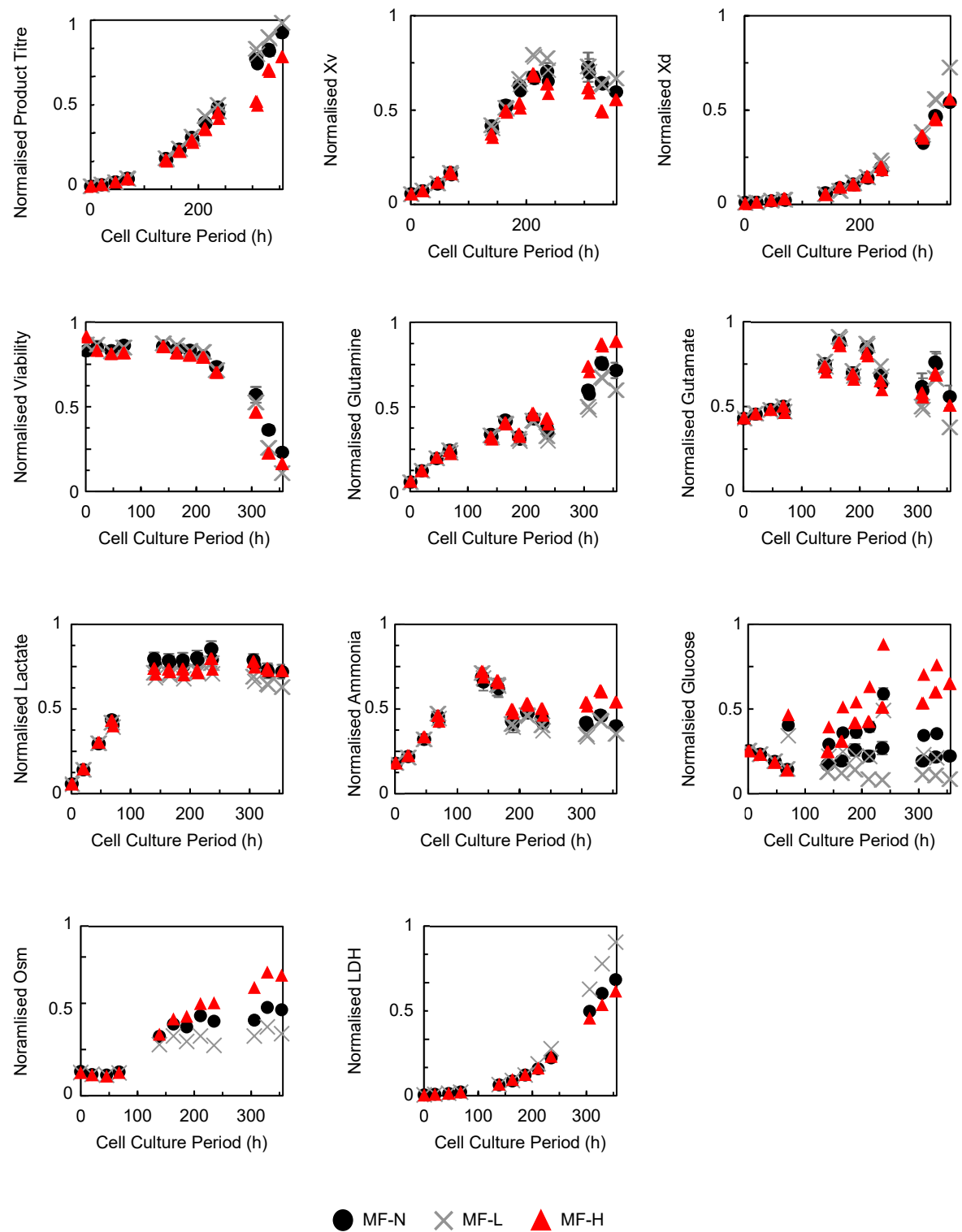


Figure 16 Effects on cell culture of a nominal glucose feeding strategy (MF-N), a low glucose feeding strategy (MF-L), and a high glucose feeding strategy (MF-H). Normalised using min-max normalisation.

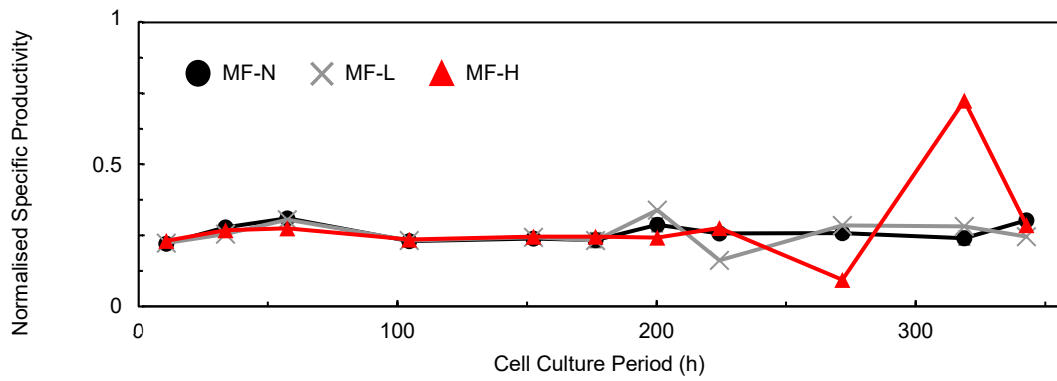


Figure 17 Specific productivity across time within the experimental model-building experiments.

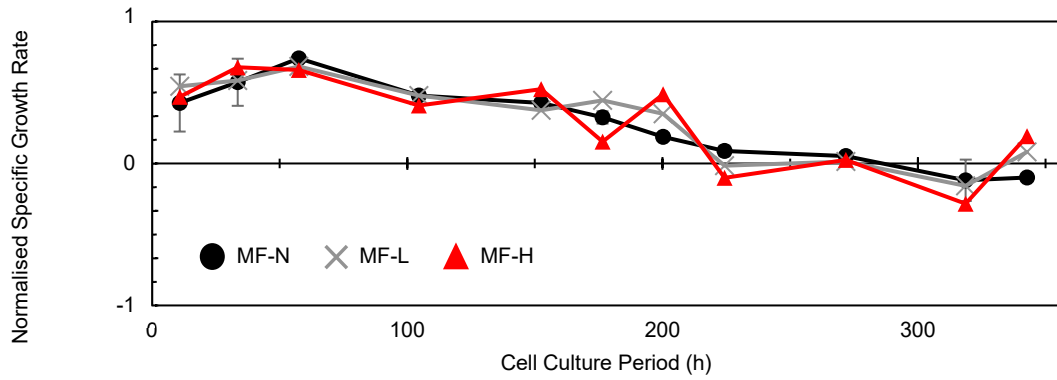


Figure 18 Specific growth rate across time within the experimental model-building experiments.

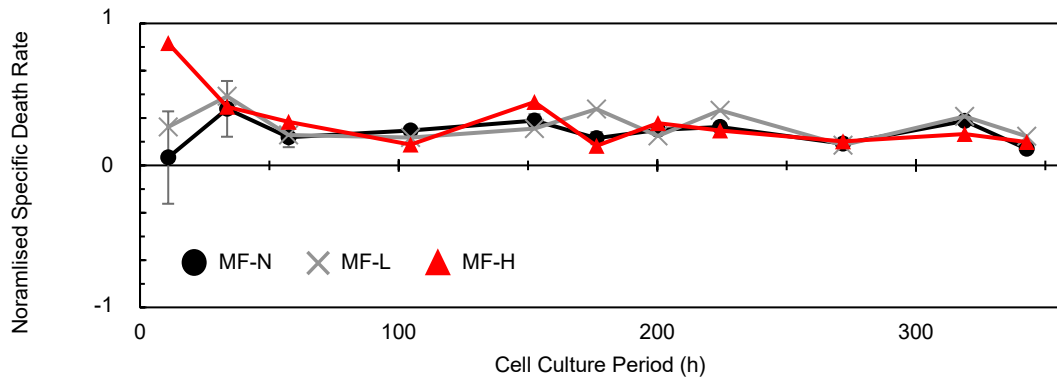


Figure 19 Specific death rate across time within the experimental model-building experiments. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a high amount of glucose. Normalised using min-max normalisation.

4.4 Model Development

The cell culture model was adapted from previous work on fed-batch CHO cell cultures (Alhuthali& Kontoravdi, 2022; Kontoravdi et al., 2010; Kotidis et al., 2019) and assumes a homogenous culture achieved through perfect mixing. It is a reduced empirical model based on the Monod growth equation. Using the rationale elaborated on below, it describes the effects of glucose, lactate, and osmolality, on cell growth and mAb productivity. The total volumetric balance is given in Eq. 10. Whilst the viable cell balance and total cell balance are given in Eq. 11 and Eq. 12 respectively.

$$\frac{dV}{dt} = F_{in} - F_{out} \quad \text{Eq. 10}$$

$$\frac{d(VX_v)}{dt} = (\mu_{growth} - \mu_{death})VX_v - F_{out}X_v \quad \text{Eq. 11}$$

$$\frac{d(VX_t)}{dt} = \mu_{growth}VX_v - F_{out}X_t - K_{Lysis}V(X_t - X_v) \quad \text{Eq. 12}$$

V (L) is the cell culture working volume, according to the volumetric flow in, F_{in} (L.h⁻¹) and out F_{out} (L.h⁻¹), X_v (cells.L⁻¹) is the viable cell density, μ_{growth} (h⁻¹) is the specific growth rate, μ_{death} (h⁻¹) is the specific death rate, X_t (h⁻¹) is the total cell density, and K_{Lysis} (h⁻¹) is the rate of cell lysis. The specific growth rate for mammalian cells can be determined by Eq. 13, a function of the limiting factor f_{lim} (-) shown in Eq. 15, and the inhibitory factor f_{inh} (-) shown in Eq. 16.

$$\mu_{growth} = \mu_{growth,max} f_{lim} f_{inh} \quad \text{Eq. 13}$$

Here, $\mu_{growth,max}$ (h⁻¹) is the maximum specific cell growth rate. The limiting factor and inhibitory factor are functions of the limiting substrates and inhibitory metabolites, respectively. The definition of these metabolites is process-dependent, as are the associated parameter values. With limited data available, it is advantageous to lessen the number of parameters in a model to avoid over-parametrisation. Hence, the selection logic of metabolites that are most influential and should be included in a Monod equation is crucial. Previously, the logic for automation of hybrid modelling has been developed (Doyle et al., 2023) where a

saturation and sigmoidal structure is used to describe cell-specific rates of metabolites. However, automating mechanistic modelling is uniquely challenging due to the requirement of maintaining a model's biological relevance using relationships unique to each metabolite instead of the same model structure across all process metabolites. Utilising previous research (Barbosa, 2021; Kontoravdi et al., 2010), in this study foundations are further demonstrated that could form the basis of future automated mechanistic modelling platforms and be subsequently applied in real-time.

To begin, experimental data is processed using the automated processing function described in Section 4.2.4 . Next data is segmented based on the growth stage, namely lag, log, stationary, and death phase. This can be done visually but, in the interest of moving towards process automation, is formalised as follows. The lag phase was split to prior the first point of inflexion in the growth rate in Figure 21 or at the time step before the first instance of $d\mu_{growth}/dt = -d\mu_{growth}/dt$ and $\mu_{growth,max}$. The log phase was identified after the lag phase where $\mu_{growth} > \mu_{growth}^{stationary,max}$, and is the maximum growth rate occurring within the stationary phase. Similarly, the stationary phase was determined after the log phase and up to where $\mu_{growth} \geq \mu_{growth}^{stationary,min}$ and is the maximum growth rate within the stationary phase. The death phase occurs after the stationary phase.

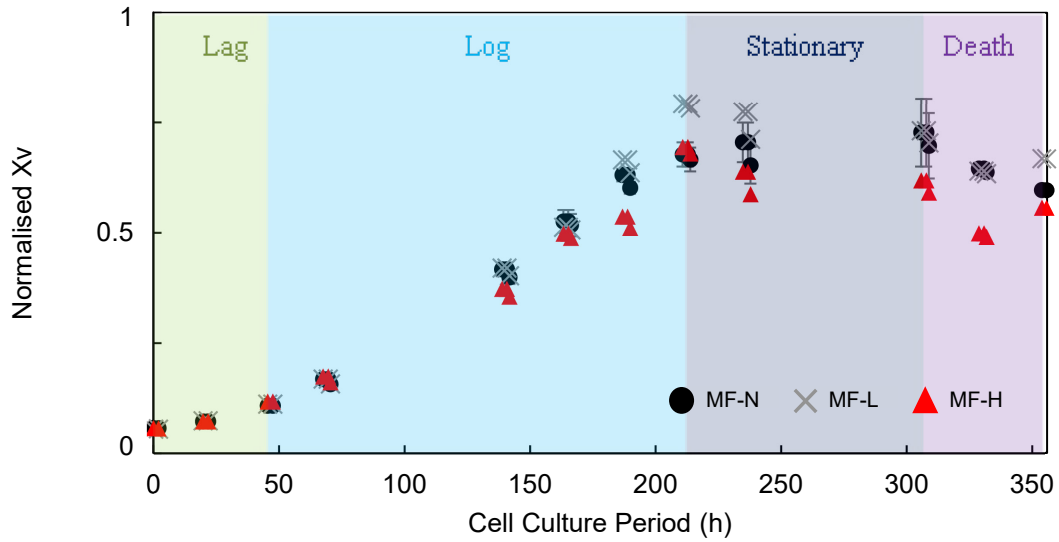


Figure 20 Viable cell density across time within the experimental model-building experiments: MF-N is with nominal glucose feeding, MF-L is with low glucose feeding, and MF-H is with high glucose feeding. Segmentation based on cell growth stage is colour-coded. Normalised using min-max normalisation.

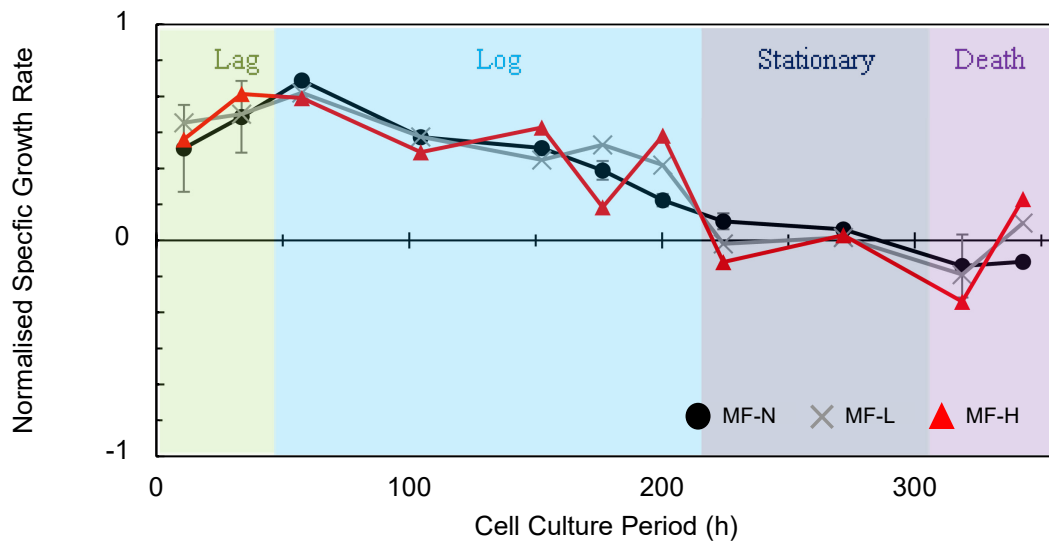


Figure 21 Specific growth rate across time within the experimental model-building experiments: MF-N is with nominal glucose feeding, MF-L is with low glucose feeding, and MF-H is with high glucose feeding. Segmentation based on cell growth stage is colour-coded. Normalised using min-max normalisation.

To determine which variables affected the growth, only experimental data occurring within the log (growth) phase of the cell culture was used to postulate a growth rate equation. Although

this approach is simplistic in that it assumes the limiting and inhibitory metabolites selected and their respective parameters remain fixed throughout all stages of the process, it is also inappropriate to burden the model due to overparameterization. Increased amounts of data may allow for a more advanced approach. Linear regression was used to find the relationships between the metabolite utilisation rates and μ_{growth} fitting Eq. 14 along with R^2 to measure fit performance. Pearson correlation coefficients (PCC) were also calculated between $\mu_{growth,predicted}^{log}$ and q_i^{log} . The results of the linear regression can be visualised in Figure 22. The R^2 , α values, and PCC were used with rules-based logic to define minor and major metabolites. As such, this will enable future automation of this workflow. In this study, as the number of experiments is small and an identical model is desired for all conditions within, the overall R^2 , α values, and PCC values were taken. With enough data and, if appropriate, one could create a set of models for varying process conditions.

The higher the value of α , the higher the predicted production of a metabolite during an increased specific cell growth rate, whilst the lower the value of α the higher the predicted consumption during increased growth. The major metabolites were set so that only metabolites with $\alpha \leq -1 \times 10^{-6}$ were classed as limiting metabolites or metabolites with $\alpha \geq 1 \times 10^{-6}$ were classed as inhibitory metabolites. Care should be taken to ensure that this decision is biologically relevant. R^2 is used to validate the linear regression and ideally should be close to 1. In this case, due to limited data, this was relaxed to reject any linear regression with R^2 less than 0.7. To further justify any relationships, especially when faced with limited data, PCC can be used. This assumes the relationship is linear during the growth phase where this analysis occurs. Inhibitory metabolites will have a PCC approaching 1, whilst limiting metabolites will have a PCC approaching -1. Overall, constraints on these metrics can be used to reject or accept metabolites into the growth equation. Tight constraints on these metrics will result in fewer metabolites included in the growth rate equation and less risk of over-parametrisation. However, this may result in the model failing to capture the full dynamics. In the latter case, an iterative approach should be taken using Figure 15 to relax the selection criteria.

$$q_i^{log} = \alpha \mu_{growth,predicted}^{log} + \beta : \mu_{growth}^{log} \subset \mu_{growth}, q_i^{log} \subset q_i \quad \text{Eq. 14}$$

Table 10 Quantitative metrics defining the relationship between growth rate and metabolite utilisation rates of minor metabolites.

	Gln			Glu			Amm		
	α	R^2	PCC	α	R^2	PCC	α	R^2	PCC
MF-N	2×10^{-7}	0.77	0.88	3×10^{-7}	0.12	0.37	1×10^{-6}	0.98	0.99
MF-L	2×10^{-7}	0.37	0.61	-3×10^{-8}	0.00	-0.02	1×10^{-6}	0.83	0.91
MF-H	1×10^{-7}	0.98	0.99	4×10^{-7}	0.52	0.72	6×10^{-7}	0.66	0.81
Overall	1×10^{-7}	0.65	0.8	3×10^{-7}	0.16	0.4	9×10^{-7}	0.73	0.85

Table 11 Quantitative metrics defining the relationship between growth rate and metabolite utilisation rates of major metabolites.

	Lac			Glc		
	α	R^2	PCC	α	R^2	PCC
MF-N	1×10^{-5}	0.94	0.97	-6×10^{-6}	0.99	-0.99
MF-L	1×10^{-5}	0.93	0.97	-6×10^{-6}	0.88	-0.94
MF-H	7×10^{-6}	0.52	0.72	-3×10^{-6}	0.53	-0.72

After assessing the relationship between metabolites utilisation rate and growth rate, the results of the analysis can be found in Table 10 and Table 11 and following this, the growth rate equation can be formulated as shown in Eq. 15 and Eq. 16.

$$f_{lim} = \frac{[Glc]}{[Glc] + K_{Glc}} \quad \text{Eq. 15}$$

$$f_{inh} = \frac{KI_{Lac}}{[Lac] + KI_{Lac}} \frac{KI_{Osm}}{Osm + KI_{Osm}} \quad \text{Eq. 16}$$

Using data analysis, glucose is considered a limiting substrate, while lactate is considered to inhibit growth. This is further clear in Figure 22 where there is a strong apparent correlation between increased glucose consumption (negative utility rate) with increased growth rate, and increased lactate production at higher growth rates. In Eq. 15 and Eq. 16 the extracellular concentrations of glucose and lactate are represented by $[Glc]$ (mM) and $[Lac]$ (mM) respectively. K_{Glc} (mM) is the Monod constant for glucose representing the concentration when glucose becomes limiting to growth. KI_{Lac} (mM) is the Monod constant for lactate representing the concentration when lactate becomes inhibitory to growth. The inclusion of lactate and glucose within the growth rate equation is also common in notable examples of existing research (Kontoravdi et al., 2010; Kotidis, Jedrzejewski, et al., 2019; Miller et al., 1988). Ammonia is often included as an inhibitory metabolite in other research (Kontoravdi et al., 2010). Further, studies suggest that the inhibitory concentration is around 2-5mM although this value is cell line and condition dependent (Altamirano et al., 2013; Xing et al., 2008). However, in the present study, ammonia was not considered as (a) data analysis showed otherwise, (b) the extracellular concentrations of ammonia were low, and (c) during parameter estimation, the value of KI_{Amm} tended to infinity.

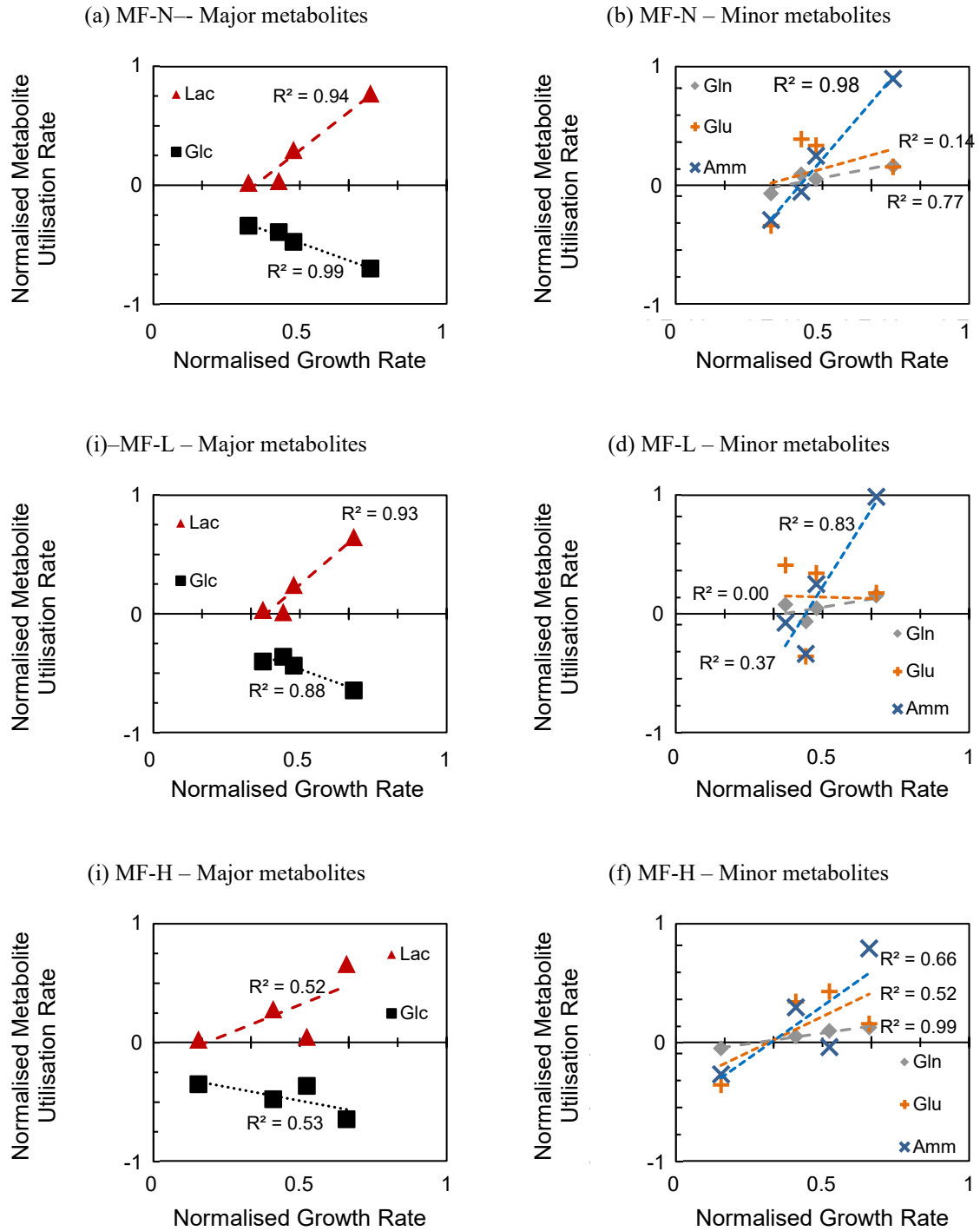


Figure 22 Relationship of growth and metabolite utilisation in (a) major metabolites in MF-N, (b) minor metabolites in MF-N, (c) major metabolites in MF-L, (d) minor metabolites in MF-L, (e) major metabolites in MF-H, (f) minor metabolites in MF-H. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a high amount of glucose. Normalised using min-max.

In Eq. 16, Osm (mOsm/kg H_2O) signifies the osmolality of the extracellular environment and KI_{Osm} (mOsm/kg H_2O) is an inhibitory constant for osmolality. The choice to include osmolality was based on the apparent effects of osmolality on growth shown in Figure 23. Here the overall specific growth rate was estimated using Eq. 6 and the overall integral osmolality was estimated using the trapezium rule. The integral osmolality is used to capture the effects over time of osmolality throughout the process, as opposed to using the endpoint value, which may not reflect the overall trajectory.

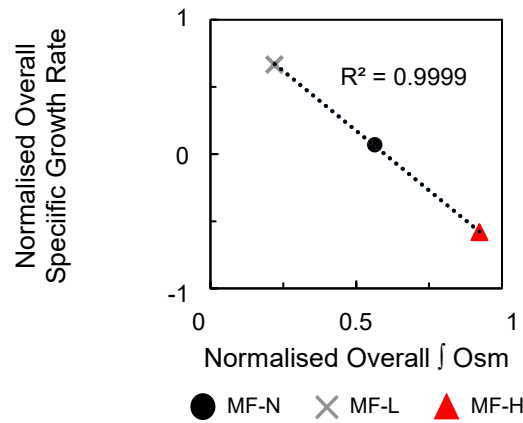


Figure 23 Effects of osmolality on specific growth rate in model-building experiments with nominal glucose feeding (MF-N), low glucose feeding (MF-L), and high glucose feeding (MF-H). Normalised using min-max normalisation.

Increased glucose feeding leads to higher osmolality of the extracellular environment due to excess glucose and potentially increased waste metabolites. This negatively affects the growth rate and is also validated within previous research (Alhuthali et al., 2021; Romanova et al., 2021; M. M. Zhu et al., 2005).

$$\mu_{death} = \mu_{death,av} \quad \text{Eq. 17}$$

Eq. 17 describes the cell-specific death rate where $\mu_{death,av}$ (h^{-1}) is the average specific cell death rate. Aside from during the lag phase of cell culture, the cell-specific death rate had a flat trend across time as shown in Figure 24. Further, towards the end of the batch, variability in measurements of viable cell density and total cell density is displayed making it harder to distinguish the overall trends in death rate. Therefore, to minimise the number of parameters

required, reducing the risk of over-parametrisation particularly due to the low amount of experimental data, the death rate was modelled as constant.

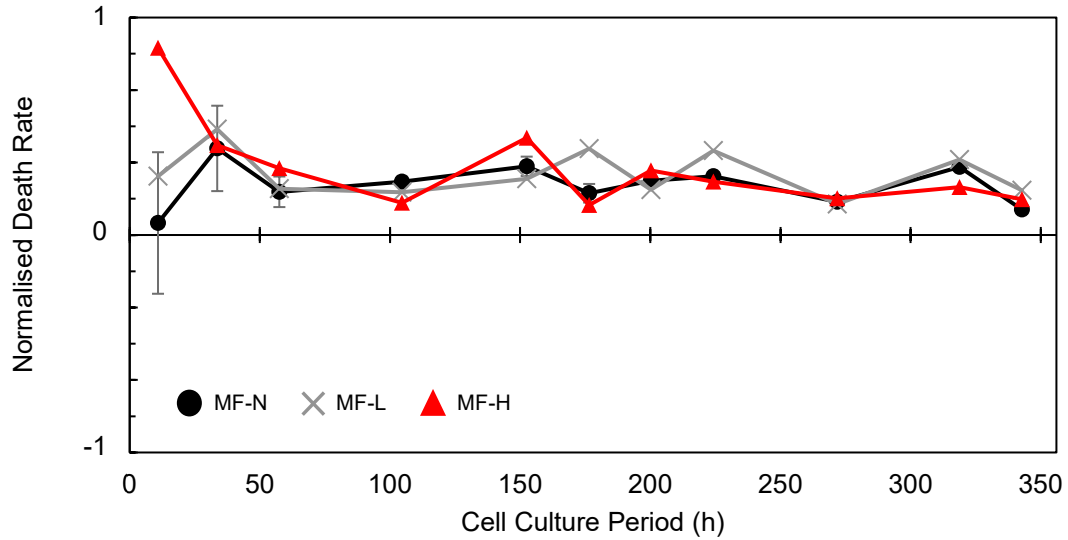


Figure 24 Death rate across the cell culture period. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a high amount of glucose. Normalised using min-max normalisation.

For real-time use, only the process variables that are manipulated, controlled, or otherwise are required to compute manipulated and controlled variables should be included within the model. This reduces the number of parameters required whilst also decreasing the number of variables that require measurement. In this case, the metabolites measured will be limited to only those in the growth rate equation, glucose, and lactate. However, other metabolites not measured in these experiments may be inhibitory or limiting to cell growth.

$$\frac{d(V[Glc])}{dt} = F_{in}[Glc] - F_{out}[Glc] + q_{glc}VX_v \quad \text{Eq. 18}$$

$$q_{glc} = -\frac{\mu_{growth}}{Y_{x_{glc}}} \quad \text{Eq. 19}$$

Eq. 18 and Eq. 19 describe the dynamics associated with the extracellular glucose concentration, $[Glc]$ (mM). Here q_{glc} (mM.cell⁻¹.hr⁻¹) is the specific utility rate of glucose and $Y_{x_{glc}}$ is the yield of cell biomass on glucose and is assumed to remain constant to avoid

overparameterization. For the same reason, the model does not include a maintenance coefficient of glucose either.

$$\frac{d(V[Lac])}{dt} = F_{in}[Lac] - F_{out}[Lac] + q_{glc}VX_v \quad \text{Eq. 20}$$

$$q_{Lac} = -\frac{q_{glc}}{Y_{LacGlc}} - m_{Lac,1} \frac{[Lac]}{Lac_{max} + [lac]} + m_{Lac,2} \left(-\frac{Osm}{K_{OsmLac}} + \alpha_{Lac} \right) \quad \text{Eq. 21}$$

Eq. 20 and Eq. 21 describe the dynamics associated with the extracellular lactate concentration, $[Lac]$ (mM). Here q_{Lac} (mM.cell⁻¹.hr⁻¹) is the specific utilisation rate of lactate which is assumed to be a function of glucose utilisation rate via the production of pyruvate in the glycolysis metabolic pathway which is subsequently converted to lactate. Y_{LacGlc} (mM_{Lac}.mM_{Glc}) is the yield of lactate from glucose. Lac_{max} (mM) is a kinetic constant for lactate consumption that occurs in the stationary phase (Hartley et al., 2018; Sou et al., 2017). When $[Lac] > Lac_{max}$ then $-m_{Lac,1} \frac{[Lac]}{Lac_{max} + [lac]}$ will approach $-m_{Lac,1}$ indicating lactate consumption. Conversely, when $[Lac] < Lac_{max}$, then $-m_{Lac,1} \frac{[Lac]}{Lac_{max} + [lac]}$ will approach 0 and the rate will be higher indicating lactate production. During lactate consumption, lactate may be converted to pyruvate which subsequently can enter the central metabolism. With limited measurements during experiments, the effects of these reactions may go unmeasured and are not described in this model. As such, other metabolic pathways from lactate can be approximated through a lactate maintenance coefficient, $m_{Lac,1}$ (mM.cell⁻¹.hr⁻¹). It can be observed in Figure 16 that the endpoint extracellular lactate consumption of MF-L is 14.6% and 15.6% lower than MF-N and MF-H, respectively. This could suggest a slight increase in lactate consumption in experiments where osmolality was lowest. To capture this in the model, a third set of terms based on osmolality were implemented into the prediction q_{Lac} based on the effects of osmolality in previous studies (Alhuthali & Kontoravdi, 2022). Here, $m_{Lac,2}$ (mM.cell⁻¹.hr⁻¹) is a term to account for the increased lactate consumption at lower osmolality dependant on K_{OsmLac} (mOsm/kg H₂O) and α_{Lac} (mOsm/kg H₂O).

$$Osm = \alpha_{glc,osm}[Glc] + \alpha_{lac,osm}[Lac] + \alpha_{fin,osm} \int F_{in} + \alpha_{osm} \quad \text{Eq. 22}$$

Osmolality was predicted using Eq. 22 and is a function of the glucose that is fed into the cell culture. $\alpha_{glc,osm}$ (mOsm/kg H₂O) and $\alpha_{lac,osm}$ (mOsm/kg H₂O) are linked to the increase in glucose and lactate respectively with osmolality. The effects of feed additions on osmolality are approximated by taking the integral of the feed in and using $\alpha_{fin,osm}$ (mOsm/kg H₂O) as a parameter to correlate the effects of feed in on osmolality. Finally, α_{osm} (mOsm/kg H₂O) is used to describe the addition of any other inputs on osmolality such as base additions where those states are not included in the model.

$$\frac{d(V[mAb])}{dt} = -F_{out}[mAb] + q_{mAb}VX_v \quad \text{Eq. 23}$$

$$q_{mAb} = m_{mAb} \quad \text{Eq. 24}$$

Eq. 23 and Eq. 24 describe the product mass balance where [mAb] (mg.L⁻¹) is the product titre and m_{mAb} (mg.cell⁻¹h⁻¹) is a non-growth associated term. In experimental studies, the specific productivity appeared independent of growth rate and constant between experiments (see Figure 17). Therefore, to reduce the risk of overparameterization, only a non-growth-associated term was included in the model.

4.5 Model Parameterisation

Using a warm start with parameters taken from previous research or data analysis, the model was parameterised (using the methodology stated in Section 4.4) with the results shown in Table 12 and Table 13. The parameters were found sequentially by fixing other variables as polynomials to reduce the degrees of freedom during parametrisation. Following this, improved guesses of parameters were found, and all variables were unfixed. The 95% confidence interval is commonly used in other studies to determine whether the parameters can be accepted if the former is smaller than 10% of the parameter value (Kontoravdi et al., 2010). In this case, the maximum growth rate, and the lysis rate had 95% confidence intervals outside of this. However, this can be justified given the limited amount of data available to parametrise the model initially. Interestingly, the yield of lactate from glucose ($Y_{Lac_{Glc}}$) was close to its theoretical maximum of 2.00. The equation for lactate utilisation rate (Eq. 20) only predicts the conversion of glucose to lactate and lactate consumption. However, there are other metabolic pathways not described here which may influence the specific lactate utilisation rate. For example, research reports how increased glutaminolysis can replenish pyruvate in the cell, which could lead to increased lactate production (Hartley et al., 2018). This could be why in Figure 16, both extracellular lactate concentration and extracellular glutamine concentration are correlated ($PCC = 0.94$) prior to the death phase. Performing flux balance analysis may provide further insight, although this is out of the scope of this research and would require significantly more measurements of metabolite concentrations which were not available in these experiments. Overall, the model is only an approximation of what is occurring biologically, and parameters (e.g. Monod limiting/inhibitory constants) are also only estimates of their true biological value. In practice, if more data became available, re-parametrisation or restructuring may enable better estimates. This may be required especially if new operating conditions were used relative to the data which the model was originally parameterised upon.

Table 12 Parametrisation results for the parameters associated with growth and productivity in the mechanistic model. *95% confidence interval is greater than 10% of the parameter value.

Estimated Parameter	Units	Confidence Interval /Value (%)
$\mu_{growth,max}$	h^{-1}	35.67*
K_{Glc}	mM	4.72
KI_{Lac}	mM	4.12
KI_{Osm}	mOsm.kg H ₂ O	0.27
$\mu_{death,max}$	h^{-1}	0.32
K_{Lysis}	h^{-1}	286.50*
m_{mAb}	mg.h ⁻¹	1.15

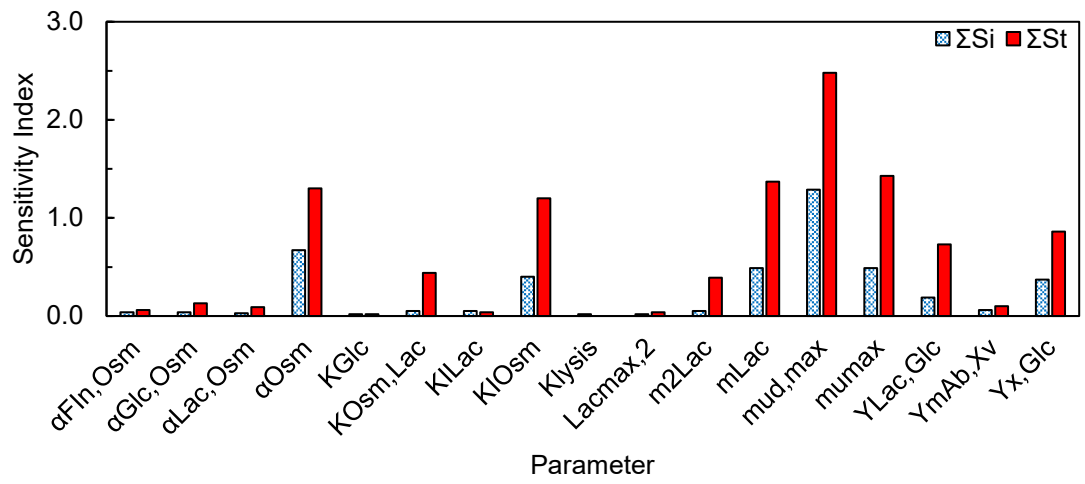
Table 13 Parametrisation results for the parameters associated with extracellular metabolite concentration and osmolality in the mechanistic model.

Estimated Parameter	Units	Confidence Interval/Value (%)
$Y_{x,glc}$	mM.cell ⁻¹ .hr ⁻¹	0.19
Y_{LacGlc}	mM.cell ⁻¹ .hr ⁻¹	8.69
$m_{Lac,1}$	mM.cell ⁻¹ .hr ⁻¹	2.66
Lac_{max}	mM	0.82
$m_{Lac,2}$	mM.cell ⁻¹ .hr ⁻¹	0.82
K_{OsmLac}	mOsm.kg H ₂ O	0.34
$\alpha_{Lac,Osm}$	mOsm.kg H ₂ O	1.39
$\alpha_{Glc,Osm}$	mOsm.kg H ₂ O	3.93
$\alpha_{Fin,Osm}$	mOsm.kg H ₂ O	7.94
α_{Osm}	mOsm.kg H ₂ O	1.18

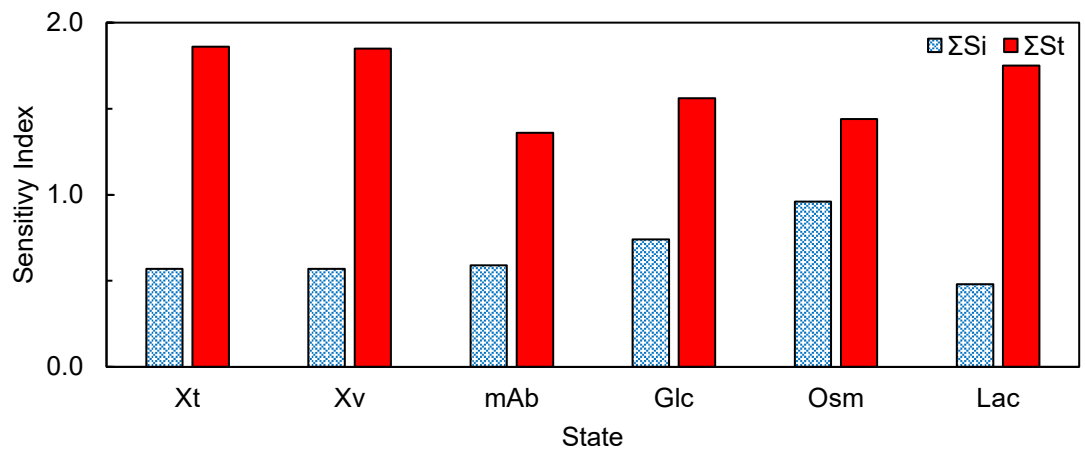
4.6 Results

4.6.1 Model Parameter Sensitivity Analysis

Sensitivity analysis of the effects of the parameters on the model outputs was carried out using the Global Systems Analysis function on gPROMS ModelBuilder 7.2 (Siemens PSE) as described in Section 4.2.4. Considering the 95% confidence intervals of the model parameters, quasi-random Sobol sampling was used to vary parameters around their value by +/- 40% as all parameters, apart from K_{Lysis} , are below this threshold. The result provides sensitivity indices that quantify the first order interactions of each parameter with the output (S_i) and the total effect the parameters have on outputs (S_t) (Saltelli et al., 2010). The first order effects can be used for factor prioritisation. In addition, the total effects include first order effects and higher order effects due to interaction between parameters and be used for factor fixing. Figure 25a shows the effects of each parameter on the model variability. Here, the first order and total order indices (S_i and S_t respectively) associated with the effect of each parameter on each of the six outputs in the mechanistic model described in Section 4.4 and Section 4.5 were summed across all outputs for each parameter to get ΣS_i and ΣS_t respectively. Further, Figure 25b shows which output is most affected directly by model uncertainty stemming from parameters. Here, the first order and total order indices (S_i and S_t respectively) associated with the effect of each parameter on each output in the mechanistic model described in Section 4.4 and Section 4.5 were summed across all parameters on each output to get ΣS_i and ΣS_t respectively. For example, in Figure 25b, the sum of the total effects on outputs is lowest in product titre suggesting that parameters associated with this affect model variability less than those associated with other parameters. A sensitivity threshold can be set to filter parameters that have a negligible contribution to model variability usually between 0-0.2 (Kiparissides et al., 2009). Although this threshold is arbitrary, in this study it was set at 0.05 so as not to class parameters as insignificant when modelling using limited experimental data (Kotidis, Demis, et al., 2019; Saltelli et al., 2004).



(a)

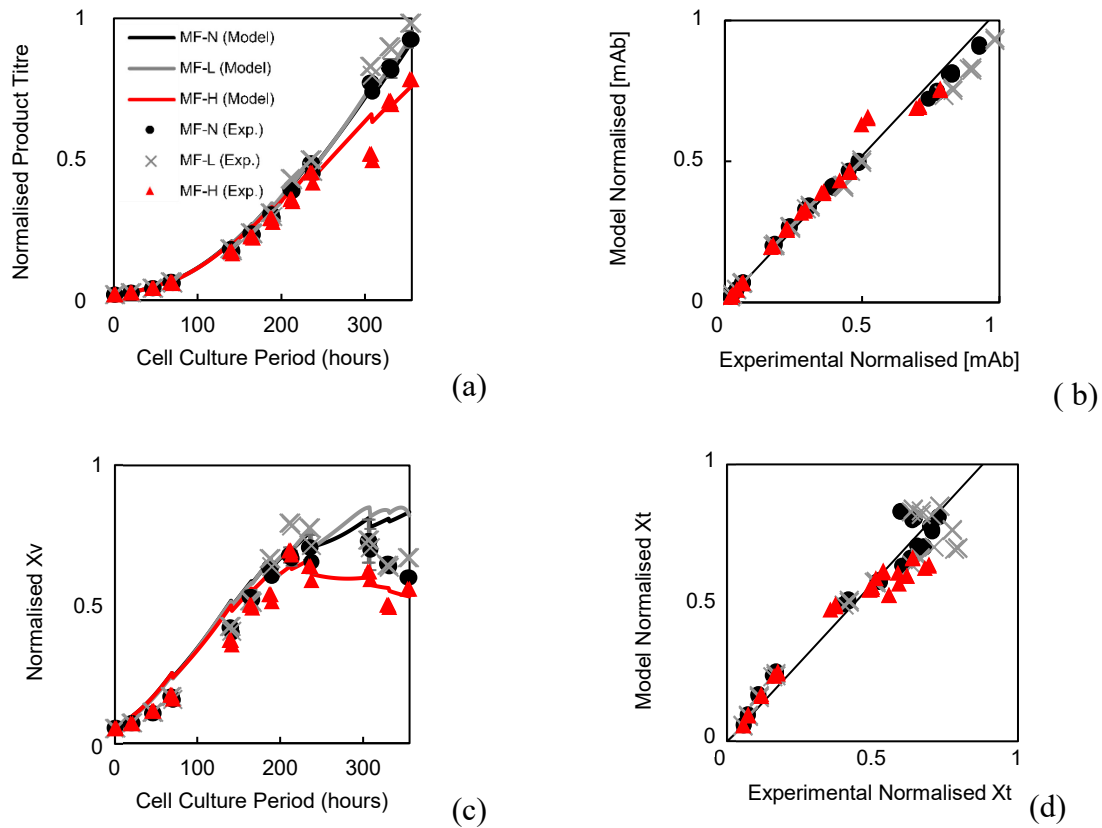


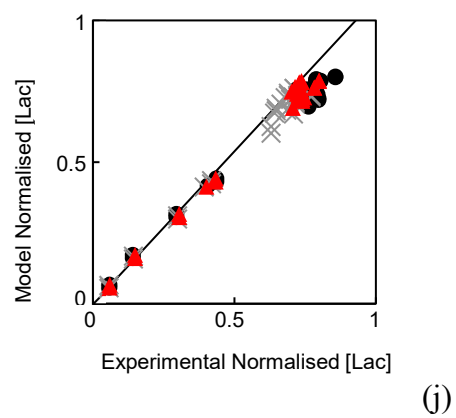
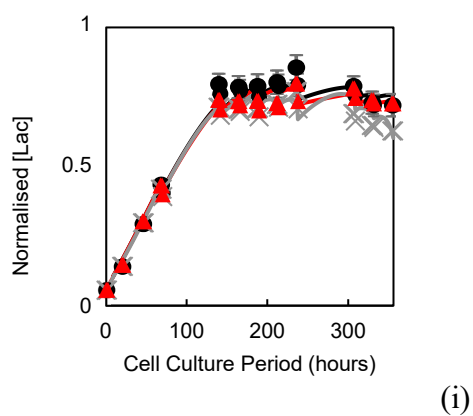
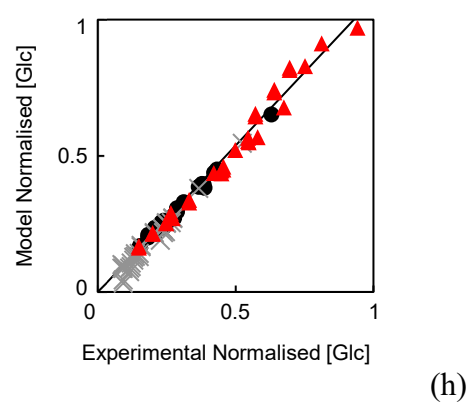
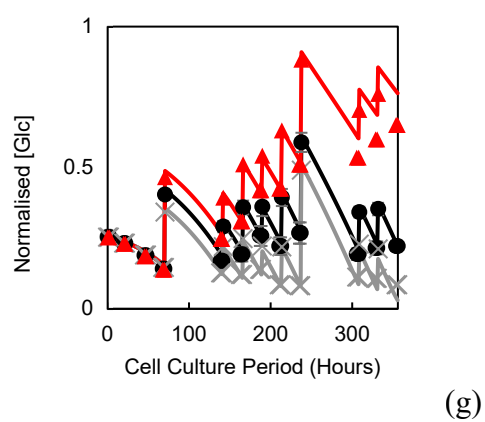
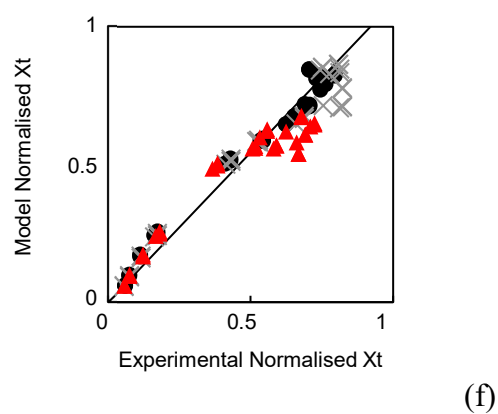
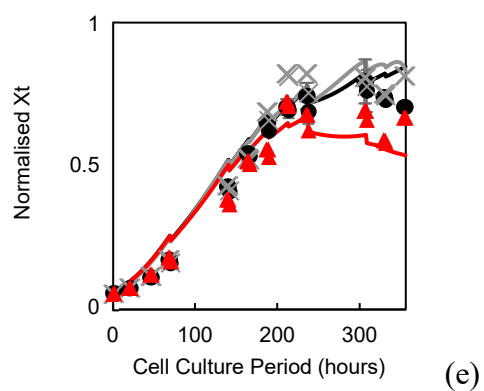
(b)

Figure 25 First order indices and total sensitivity indices from the parameter sensitivity analysis varying parameters by 40% on (a) summing indices on each parameter, and (b) summing indices on each output.

4.7 Model Performance

Using experimentally validated mathematical structures, a biologically relevant model is postulated that estimates the specific cell growth rate, the specific cell metabolite utilisation rates, and the specific cell productivity. In doing so, the extracellular concentration of metabolites, cell density, and product titre can be predicted using only the initial viable cell density, the feeding schedule of the fed-batch experiments, and the metabolite concentrations in the feed and media. Figure 26 shows the results of the model's predictive capability when compared to experimental data across the cell culture period. Here, parity plots are shown where a perfect fit is $y=x$.





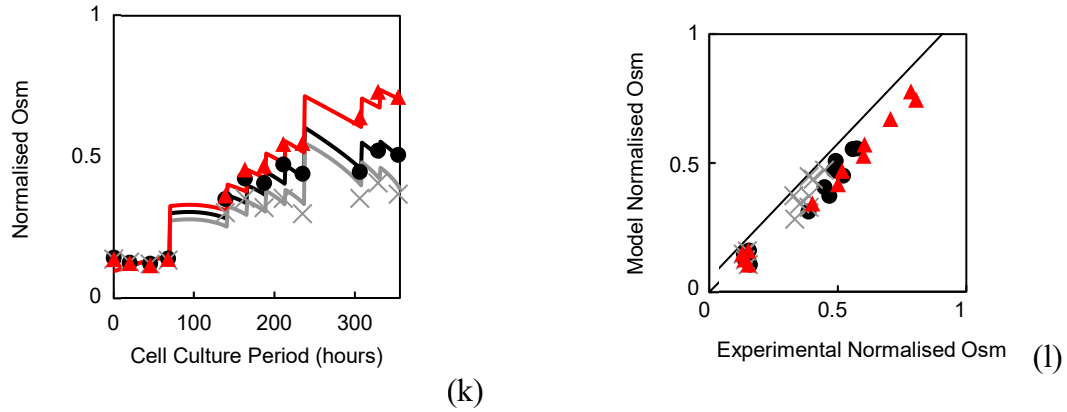


Figure 26 Experimental vs model predictions of variables within the mechanistic model across the cell culture period, and in parity. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose. Normalised using min-max normalisation.

In addition, the values of R^2 , RMSE, and Pearson correlation coefficient (PCC) are also used to score the goodness of fit between the model and experimental data as shown in Table 14. Overall, Table 6 shows that the model performed well achieving accurate predictions of the experimental data with all R^2 being above 0.92 and PCC above 0.96. Under this metric, this model is on par with existing mechanistic models presented in literature (Bayer et al., 2023). The model also performs well when looking at the trajectories and parity plots in Figure 14.

Table 14 Performance metrics for the goodness of fit between model predictions and experimental data. RMSE: Root mean square error, PCC: Pearson Correlation Coefficient. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose.

	MF-N			MF-L			MF-H		
	R^2	RMSE	PCC	R^2	RMSE	PCC	R^2	RMSE	PCC
[mAb]	1.00	62.51	1.00	1.00	108.98	1.00	0.97	142.00	0.99
Xv	0.96	1.89×10^6	0.98	0.92	1.90×10^6	0.96	0.95	1.22×10^6	0.97
Xt	0.98	1.25×10^6	0.99	0.96	1.31×10^6	0.98	0.92	1.40×10^6	0.96
[Glc]	0.99	2.37	1.00	0.97	3.10	0.99	0.98	7.63	0.99
[Lac]	0.99	1.98	0.99	0.99	1.73	0.99	0.99	1.50	1.00
Osm	0.96	11.37	0.98	0.92	9.26	0.96	0.99	11.65	0.99

4.8 Discussion

Viable cell density is accurately described during the lag, log, and stationary stages of cell culture. However, during the death phase, the model over-predicts the viable cell density of MF-N and MF-L (Figure 26). The assumption that cell death can be described by one parameter as shown in Eq. 17 may contribute to an inaccurate description of viable cell density. Despite, the confidence interval obtained during parametrisation being within an accepted 10%, the sensitivity analysis revealed that the model is highly sensitive to this parameter. As such a poor approximation of cell death will affect model accuracy negatively. In addition to this, looking at the decrease in both viable cell density and total cell density that occurs during the death phase (Figure 26), there seems to be a degree of cell lysis that may not be captured by the model. This risk is further shown during the parameter sensitivity analysis wherein the confidence interval of K_{Lysis} is high. Here, a weak parameter estimate due to limited data may be the cause of the suboptimal description of viable and total cell density during the death phase. Although the majority of product is made during the lag, log and stationary phases, and a batch can be stopped prior to the death phase, poor predictions of the late-stage culture, can lead to increased amounts of undesirable impurities including host cell proteins and DNA released by cells during secretion and cell lysis. In general, there is less mechanistic understanding of cell death (versus other growth stages) and even recent models utilising a hybrid approach to describe cell growth and death dynamics achieved similar R^2 values to those in the present study (Okamura et al., 2022). The total cell density (Figure 26) displays similar characteristics to the viable cell density and, during the death stage of cell culture, the model appears to predict experimental results more closely. That said, the challenges and risks associated with parameterisation during instances of low data are evident in Figure 13. Here, parameters within the growth rate (Table 4) equation including $\mu_{growth,max}$ and KI_{Osm} have relatively large sensitivity indices indicating that any inaccuracies in the predicted values of these parameters could cause a flawed prediction of the growth rate and viable cell density. Further, $\mu_{growth,max}$ has a large confidence interval indicating a high degree of uncertainty in its predicted value. However, the drawbacks and risks mentioned formerly are outweighed given the model's overall performance in the prediction of viable cell density and total cell

density. Importantly, the model correctly predicts the stunted growth during increased osmolality especially during the MF-H experiment.

One of the key performance indicators is the product titre, its dynamic profile and its endpoint value. Here the model displayed excellent performance as indicated by the PCC value. Further, the RMSE for all experiments apart from MF-H was approaching the experimental error (up to 80 mg.L⁻¹). During MF-H, there may have been a measurement error on Day 13 evident in Figure 14 which contributed to the slightly higher RMSE than in other experiments. However, by the end of the batch, this was rectified, and the model had accurately predicted the endpoint product titre, especially during increased osmolality in MF-H. Promisingly, the sensitivity analysis shown in Figure 13 showcased the model's insensitivity to parameters associated with product titre shown by first order effects indices. In particular, the sum of the total effects indices on product titre is the lowest out of all other modelled variables. This could be put down also because this variable presented limited non-linear behaviour across time relative to other variables. This is ideal given that product titre is a key performance indicator of the process and in future real-time applications of the model, the ability to predict this with certainty will be crucial for robust control.

Next looking at the manipulated variable, extracellular glucose concentration, the model correctly predicts the consumption of glucose. This is indicated in Figure 14 with a PCC of over 0.99 in all cases shown in Table 6. However, Figure 14 shows some discrepancies in predicting extracellular glucose concentration during late-stage culture in MF-H. This could be due to inaccuracy in the prediction of viable cell density. Yet, viable cell density is predicted most accurately during this stage in MF-H compared to other experiments. As such, it is plausible that the cell-specific glucose utilisation rate may be inaccurately modelled. Interestingly, the predictions of extracellular glucose concentration for MF-N and MF-L experiments have a lower RMSE compared to MF-H as shown in Table 6. Yet in both experiments, the viable cell density is overpredicted. This suggests that the true specific glucose consumption rate is lower than predicted during these MF-N and MF-L despite yielding more accurate predictions of the absolute extracellular concentration of glucose. Despite this, the R^2 of all model predictions for extracellular glucose concentration is above 0.95 overbearing any minor inadequacies.

Figure 14 shows that the model accurately predicts the lactate metabolic shift, a key behaviour of productive cell lines. Although there are few differences during the cell culture between the experiments, the model can capture that MF-L has a lower extracellular lactate concentration and displayed signs of increased lactate consumption towards the end of the cell culture. Crucially, the R^2 and PCC all are above 0.99 indicating an excellent model fit.

Finally, the model can accurately estimate osmolality despite only predicting the extracellular concentration of two metabolites. This is done through an algebraic approach inspired by previous studies (Alhuthali & Kontoravdi, 2022). In doing so, the variation in osmolality and subsequent effects on viable cell density are accurately predicted.

The model presented herein is well suited for real-time application in that process monitoring requirements are minimised due to low reliance on complex offline measurement technologies. For example, to perform real-time model-based control, efforts should be minimal to update the controller with the most recent measurements. The future vision is to combine the methodology described in this chapter towards automated mechanistic modelling platforms for real-time application. Here, a model can be used as a digital shadow where predictions can be made one step ahead, or as a digital twin wherein a mechanistic model derived prior to process operation can be deployed and the constituent equations and parameters updated upon receipt of new data.

4.9 Summary

In Chapter 4, a set of experiments in 2L reactors was used to understand the effects of glucose feeding on a cell culture. Previously reported results at a small scale were validated where reduced growth was observed at increased osmolality. This highlights the risk of overfeeding. In addition, a mechanistic model was developed outlining existing modelling frameworks (Kontoravdi et al., 2010) here with the aim of producing a model during instances of limited data availability. The mechanistic model accurately models the cell growth, and cell culture productivity and predicts the extracellular metabolite concentrations of those that most impact growth. Importantly for real-time application of this model expected in further research, the model is limited to a minimal number of metabolites that do not require complex measurement techniques. As such, the model can be deployed at ease in real-time without hindrance to efficient operation and the included metabolites can be monitored or controlled to ensure optimum process performance.

Attention was given towards automating the data processing from raw experimental data to be used further during model-building. This is also the first step towards developing further automation tools. In addition, the logic towards minimising the metabolites required to accurately model the process was outlined to balance between over-parametrisation during limited data availability, and a model's ability to explain core mechanisms such as growth or productivity. This involved segregation of the process into key growth phases namely lag, log, stationary, and death. Next, a detailed analysis of each measured metabolite's relationship with cell growth was performed. In this instance, only glucose and lactate were required to explain the cell culture growth. Finally, osmolality was incorporated into the growth rate equation inspired by previous literature albeit in this study a reduced-size model was built considering the data availability challenges.

Overall, the model was able to describe core mechanisms accurately including predicting growth and productivity. In addition, key metabolites were described correctly, and the model was able to correctly predict key mechanisms including a metabolic shift from lactate production to consumption. The model can describe the effect of variations in glucose feeding to the cell culture and is specific to the cell line and scale.

Chapter 5 Advanced Process Control of CHO Cell Cultures

5.1 Introduction

Despite the large market growth in the use of monoclonal antibodies as therapies for major diseases, the great cost of biologics, and the development of biosimilars, have created an increased focus on developing lower-cost manufacturing methods (Reddy et al., 2023; G. Walsh & Walsh, 2022). This includes the use of process automation which can reduce cost per unit of product, increase process performance, increase process robustness, or shorten process development timelines (Luo, Kurian, et al., 2021; Mitra & Murthy, 2022). However, implementing process automation requires the deployment of advanced process control, this remains a formidable challenge (Aehle et al., 2012). Here, there are several aspects contributing to a lack of advanced process control in CHO cell cultures including (1) the high degree of non-linearity in the process, (2) a lack of robust and inexpensive process sensor technologies, (3) a weak process understanding, (4) greatly interdependent process variables, (5) a highly multivariable process, and (6) the presence of large time-delays in dynamics (Rathore et al., 2021). Thus, significant research gaps are present as existing process control of mammalian cell cultures is basic (e.g., regulatory control of (pH, dissolved oxygen, and temperature) compared to the advanced control present in other process industries (Mitra & Murthy, 2022). Addressing these research gaps will also tackle one of the final hurdles towards wholly implementing quality by design (Anurag S Rathore & Helen Winkle, 2009).

5.1.1 Operating Strategy Development

Process control firstly requires a defined operating strategy subsequently implemented by a defined process control algorithm. As a basic example, process control of temperature in a CHO cell culture would entail two decisions, namely the choice of temperature, e.g., 35°C, and the choice of process control algorithm e.g., Proportional-Integral-Derivative (PID). Here, an operating strategy refers to a predefined objective and could include the following: (1) a fixed setpoint value such as controlling the cell culture temperature to 35°C, (2) a fixed setpoint range for example, controlling the cell culture pH between 6.9-7.0, (3) a dynamic setpoint for instance, performing temperature shifts during process operation (McHugh et al., 2020) or (4) a function to be minimised or maximised, for example, maximising product titre (Naderi et al., 2014). Despite clear over-arching goals of the biopharmaceutical industry to produce products in a cost-effective, resource-minimising manner, the definition of process-specific operating strategies is non-trivial. Previously, an operating strategy would be defined through extensive experimental studies, with conditions defined by one-factor-at-a-time procedures (Bibila & Robinson, 1995). At present, design of experiments studies are common but still require large experimental efforts (Möllerl et al., 2015). In part, this perceived large experimental effort required to develop a suitable operating strategy has hampered the adoption of advanced process controllers (Downey et al., 2017). To address this, model-based decision-making using *in silico* tools is becoming increasingly utilised as for example they may reduce development costs or time by maximally utilising existing process data and allow for efficient design space analysis (Ahleboot et al., 2021; Helgers et al., 2022; Reddy et al., 2023). In recent research, such tools have been used to decipher the optimal pH shift to maximise productivity using genetic algorithm-based optimisation (Hogiri et al., 2018), tailor the feeding schedule for optimal glycan distribution using sensitivity analysis (Kotidis, Demis, et al., 2019), or to improve product titre using design of dynamic experiments (Luo, Stanton, et al., 2023). Intrinsically, the development of an advanced process control strategy should be paired with an operating strategy defined with model-based tools.

5.1.2 Open-Loop vs Closed-Loop Systems

Once an operating strategy is developed, process control may run in either an open-loop manner or a closed-loop manner. In open-loop, the control inputs are independent of the real-time process outputs, such as using dynamic optimisation to form a pre-defined feeding regime to maximise galactosylation (Kotidis, Jedrzejewski, et al., 2019) or simultaneously product titre and galactosylation (Kappatou et al., 2020). Alternatively, in closed-loop control, the control inputs are decided during process operation and are dependent on the process outputs. For example, feeding is based on a feedback loop containing the most recent set of measurements to maximise glycosylation (Luo, Kurian, et al., 2023). One of the main advantages of closed-loop control that makes it suitable for CHO cell cultures is its inherent ability to correct for disturbances and plant model mismatch (Rathore et al., 2021). There are notable examples of closed-loop control of fed-batch processes that include monitoring bioreactor conditions such as pH (Gagnon et al., 2011; Konakovsky et al., 2016), extracellular metabolite concentrations (Berry et al., 2016; Eyster et al., 2021; Lu et al., 2013; Webster et al., 2021), or by estimating oxygen transfer rate (Goldrick et al., 2018). Further, advanced closed-loop control of key performance indicators or critical quality attributes has also been demonstrated *in silico* to control product titre (Papathanasiou et al., 2017) and galactosylation (Kotidis et al., 2021; Luo, Kurian, et al., 2023). Yet there remains a gap in research for experimental demonstrations of closed-loop control of key performance indicators or critical quality attributes on fed-batch CHO cell cultures due to the challenges described in previous sections.

5.1.3 Key Process Control Algorithms

Regarding the process control algorithm used by the control loop, many other process industries frequently implement multi-input multi-output (MIMO), non-linear, advanced process control strategies predominantly using model predictive control (MPC) (Bauer & Craig, 2008b). However, over 90% of biopharmaceutical processes rely on more basic single-input single-output (SISO) classical process control strategies e.g., proportional–integral–derivative control (PID) (Mitra & Murthy, 2022). PID control is a reactive control algorithm that is easy to implement and highly suited to linear, time-invariant control problems. Hence, PID is frequently used within upstream biopharmaceutical processing for basic regulatory

SISO control of critical process parameters such as temperature, pH, or DO in segregated control loops (Rathore et al., 2021). To complement these control loops, the status quo in manufacturing is for a feeding recipe defined prior to process operation to be delivered manually to achieve control of key performance indicators and critical quality attributes in an open-loop manner.

In recent research in the control of CHO cell cultures, proportional-integral (PI) control has been presented for control of extracellular glucose concentration by manipulating glucose feed having secondary effects of improving product quality (Berry et al., 2016). Lately, Raman spectroscopy has been used alongside PI control to control glucose and lactic acid extracellular concentration leading to increased galactosylation (Eyster et al., 2021). Despite their achievements, these studies are examples of regulatory control of extracellular metabolite concentrations. In fact, the aforesaid benefits of PID control can diminish when implemented into CHO cell culture systems for closed-loop control of key performance indicators such as product titre (Sarna et al., 2023). Firstly, CHO cell culture systems have a large number of inputs and outputs which are also highly coupled systems. That said, the decoupling of large multivariable processes remains an open challenge in process control research, outside the scope of this study, and is one of the main problems cited by controller manufacturers (Garrido et al., 2016). Meanwhile, existing solutions increase in complexity with dimensionality, can achieve poor disturbance rejection, or achieve inferior controller tracking when implemented within nonlinear systems (Garrido et al., 2016; Navratil et al., 2020). Another more prominent challenge inhibiting the application of advanced process control to fed-batch CHO cell culturing is related to the non-linear and time-varying process dynamics. However, recent advances in process modelling have brought increased process understanding therefore rendering predictive model-based control methods promising for closed-loop control of fed-batch cultures (Bolmanis et al., 2023). Unlike PID, which is a reactive control method where inputs are decided based on past or current outputs, model predictive control can include future predictions to determine process inputs. Moreover, model predictive control can more easily handle MIMO processes via the inclusion of variables within an objective function paired with soft or hard constraints. It should be noted, however the effectiveness of this heavily depends on the accuracy of the process model. Nonetheless, advances in process modelling (discussed in Chapter 4) provide a strong direction to model predictive control of CHO cell cultures.

5.1.4 Developments in Model Predictive Control of CHO Cell Cultures

There have been several notable implementations of model predictive control in CHO cell cultures as shown in Table 15. This differentiates these contributions for example, based on mode of operation, controlled variable, manipulated variables, objective function type, or whether they are in-silico or experimental demonstrations.

Table 15 Summary of key contributions implementing model predictive control to CHO cell cultures. MVs: manipulated variables. CVs: controlled variable. OF: objective function. FB: fed-batch. DD: data-driven. SPEM: setpoint error minimisation. CVM: CV minimisation/maximisation

Reference	Mode of Operation	MVs	CVs	OF	Model Type	Demonstration
(Aehle et al., 2012)	FB	Fin _{Gln}	Total cumulative oxygen uptake rate	SPEM	Mechanistic	Experimental
(Craven et al., 2014)	FB	Fin _{Glc}	[Glc]	SPEM	Mechanistic	Experimental
(Jones et al., 2023)	FB	O ₂ and air sparge, CO ₂ sparge	DO, pH	SPEM	DD	Experimental
(Joynes & Zhang, 2022)	FB	Fin _{Glc}	[Glc]	SPEM	DD	<i>In-silico</i>
(Luo, Kurian, et al., 2023)	FB	Setpoints of agitation, temperature, DO, feed rate containing glucose, temperature, DO	[mAb], Viability, fucosylation	SPEM	DD	<i>In-silico</i>

(Ohkubo et al., 2023)	FB	Fin_{Glc} , Fin_{Gln}	[mAb]	CVM	Mechanistic	<i>In-silico</i>
(Papathanasiou et al., 2017)	FB	Fin_{Glc} , Fin_{Glu} , Fin_{Arg} , Fin_{Asp} , Fin_{Asn}	[mAb]	CVM	DD	<i>In-silico</i>
(Rashedi et al., 2023a)	FB	Fin_{Glc}	X_v , [Glc], [Glu], X_t , [Lac], [Amm], Osm, viability, [Gln], [K], [Na]	SPEM	DD	Experimental
(Rashid et al., 2021)	FB	(1 & 2) Feed containing glucose and glutamine	(1) [Glc], (2) [mAb]	(1) SPEM, (2) CVM	DD	<i>In-silico</i>
(Schmitt et al., 2019)	FB	pH setpoint and feed rate	[Lac]	SPEM	DD	Experimental
(Downey et al., 2017)	FB	Fin_{Gal}	Galactosylation	SPEM	DD	<i>In-silico</i>
(Sarna et al., 2023)	Perfusion	Temperature, pH, $[Glc]_{in}$, Fin_{Glc} , bleed rate	[mAb], harvest rate	CVM	DD	<i>In-silico</i>

(Zupke et al., 2015)	Perfusion	Fin _{Man}	High Mannose	SPEM	Mechanistic	Experimental
----------------------	-----------	--------------------	--------------	------	-------------	--------------

For example, model predictive control has been used in fed-batch cell cultures (Aehle et al., 2012; Craven et al., 2014; Downey et al., 2017; Jones et al., 2023; Joynes & Zhang, 2022; Luo, Kurian, et al., 2023; Ohkubo et al., 2023; Papathanasiou et al., 2017; Rashedi et al., 2023a; Rashid et al., 2021; Schmitt et al., 2019) and perfusion-based CHO cell cultures (Sarna et al., 2023; Zupke et al., 2015). Although there are significant research efforts towards continuous perfusion cell culturing (Sarna et al., 2023), the fed-batch mode of operation is widely implemented within industry (G. Walsh & Walsh, 2022) and therefore it is of initial importance to develop advanced control strategies for this predominant manufacturing method.

Model predictive control has been used for control of metabolite concentrations (Joynes & Zhang, 2022; Schmitt et al., 2019), pH (Jones et al., 2023), oxygen consumption (Aehle et al., 2012), viable cell density, or control of key performance indicators (e.g. product titre) or critical quality attributes (e.g. glycosylation) (Downey et al., 2017; Luo, Kurian, et al., 2023; Ohkubo et al., 2023; Papathanasiou et al., 2017; Rashedi et al., 2023a; Rashid et al., 2021; Sarna et al., 2023; Zupke et al., 2015). Although there is value and challenges remaining in regulatory control of base bioreactor conditions (pH, temperature, dissolved oxygen), metabolite concentration and other critical process parameters, it is more ambitious yet intuitive to provide supervisory control of key performance indicators such as product titre. Moreover, direct control of critical quality attributes would facilitate overcoming the final hurdle towards quality by design (Anurag S Rathore & Helen Winkle, 2009).

Model predictive control uses a mathematical model to predict the future behaviour of the system over a finite time horizon (prediction horizon). Based on the most recent measured or estimated value of the system states, an optimisation problem is solved iteratively to determine the values of the manipulated variables. A central decision during the design of model predictive control is the objective function which could contain a constant setpoint, time-dependant setpoint, or economic objective. The inclusion of constant setpoints in an objective function is useful for regulatory control e.g., to minimise the error against, and control pH to, a fixed value. Meanwhile, time-dependant setpoints are useful for following a time reference trajectory, e.g., to minimise the error against a predefined product titre trajectory. Finally, economic objectives are used to minimise/maximise a performance specification e.g., maximise product titre. Relevant studies have used setpoint tracking to minimise the error

against a predefined single setpoint (Craven et al., 2014; Jones et al., 2023; Joynes & Zhang, 2022; Rashedi et al., 2023a), trajectory tracking to minimise the error against a predefined trajectory (Aehle et al., 2012; Luo, Kurian, et al., 2023; Schmitt et al., 2019), or control variable minimisation/maximisation (Ohkubo et al., 2023; Papathanasiou et al., 2017; Rashid et al., 2021; Sarna et al., 2023). Setpoint tracking and trajectory tracking are aimed towards increasing process robustness and are useful for regulatory control (Schwenzer et al., 2021). Whereas, minimisation/maximisation of a control variable is directed more towards increasing process performance. Although error-minimising methods may reduce batch-to-batch variability, a weakness of error-minimising methods such as set-point tracking and trajectory tracking, is that they are inherently restrictive and undermine the key advantage of MPC: optimality (Schwenzer et al., 2021). The decision of setpoints or reference trajectories must be performed prior to operation. However, given that it is unlikely for any process to operate at its theoretical global optimum for any substantial length of time, there will always be room for improvement (Ellis et al., 2014). Further, any major disturbances may lead to a trajectory that is no longer feasible. For a highly non-linear process, there may be multiple routes towards achieving optimality. As such, to enable the greatest flexibility to always push towards optimal conditions defined in real-time based on feedback, rather than optimal conditions defined prior to process operation where disturbances may not have been considered, this study will focus on minimisation/maximisation of a control variable without error minimisation.

There are various examples of process models used for model predictive control of CHO cell cultures. There are very few examples where mechanistic models have been utilised for model predictive control (see Table 15). Here they are implemented within a non-linear model predictive control loop. Alternatively, data-driven models are more frequently utilised (see Table 15). Data-driven models often have a more standardised workflow, and can more easily be optimised directly for parsimony (e.g., using Bayesian optimisation within hyper-parameter tuning), thus their development can be embedded more easily into an automated workflow. In addition, linear data-driven models may have smaller computation times when implemented into model predictive control than mechanistic models deployed within nonlinear model predictive control (Camacho et al., 2007). However, fed-batch CHO cell cultures have slow dynamics relative to many chemical processes. As such, long computational times can be afforded to a greater degree. In addition, non-linear model predictive control is sometimes

preferred over linear MPC as it can directly treat nonlinear dynamics and constraints explicitly rather than using linear approximations (Gros et al., 2020). Overall, mechanistic models may be preferable over data-driven models for model predictive control applications as (a) mechanistic models can be built with increased accuracy in instances of low data availability (Bayer et al., 2023), and (b) mechanistic models can extrapolate to a greater degree relative to data-driven models (Bayer et al., 2023). Indeed, if the latter is true, then during instances where a model is built with limited data and there are process deviations where a model may be extrapolating outside of the conditions that it was built upon (e.g. if pH falls outside of the setpoint range due to increased lactate production), then mechanistic models may be advantageous here versus data-driven models. Here, a mechanistic model may be able to make more accurate predictions, and thus lead to increased controller robustness. Considering these reasons, it is of interest to experimentally demonstrate the performance of a non-linear model predictive controller that utilises an existing mechanistic model.

Finally as shown in Table 15, there is a mix of *in silico* demonstrations of model predictive control (Downey et al., 2017; Joynes & Zhang, 2022; Luo, Kurian, et al., 2023; Ohkubo et al., 2023; Papathanasiou et al., 2017; Rashid et al., 2021), and experimental demonstrations (Aehle et al., 2012; Craven et al., 2014; Jones et al., 2023; Rashedi et al., 2023a; Sarna et al., 2023; Schmitt et al., 2019; Zupke et al., 2015). Despite significant advances in advanced process control strategies for fed-batch CHO cell cultures demonstrated *in silico*, there remains a gap to subsequently deploy and compare these experimentally. This is partially a consequence of the lack of reliable process analytical technologies. Of the above experimental studies, some implement PAT online alongside model predictive control including Raman (Craven et al., 2014) and quadrupole mass spectrometry (Aehle et al., 2012). However, such technologies are not yet standard across mammalian cell culture processes which instead rely frequently on offline analysis (Chopda et al., 2022). Therefore, there remains a case to showcase advanced process control using widely available process sensor technologies.

5.1.5 Summary

To summarise, there are several major research gaps outlined below before model process control can be widely implemented in CHO cell cultures:

- There remains significant room for further research applying model predictive control to either fed-batch or perfusion operation modes. This study will focus on the dominant platform, fed-batch cell cultures.
- There are notable achievements to date in model predictive control being proposed for the control of a variety of bioreactor conditions during fed-batch CHO cell culturing. However, there is a clear lack of examples of supervisory control of key performance indicators or critical quality attributes.
- Previous research in the modelling of fed-batch CHO cell culturing has presented the advantages of mechanistic models. Yet there remains a very limited number of examples where the promise of mechanistic models is showcased during model predictive control.
- There remains an open challenge given that, to the best of the author's knowledge, there are no experimental demonstrations in literature where mechanistic models are utilised within model predictive control to control product titre.

This research will focus on the development of an advanced process controller for fed-batch CHO cell culturing that directly tracks multiple variables including product titre, utilising a mechanistic model, whilst only relying on widely available process sensor technology. This requires the efficient development of an operating strategy and subsequent implementation of a suitable closed-loop process control algorithm.

Overall, Chapter 5 aims to investigate and provide insight into the following:

- Can non-linear model predictive control be demonstrated experimentally, using a mechanistic model built with limited experimental data, to direct a CHO cell culture to optimum operating conditions?
- How can an operating strategy, including an objective function, constraints, and choice of manipulated and controlled variables be defined using *in silico* tools?
- What are the limits to which non-linear model predictive control can be implemented and how does it perform during a severe operator error?

5.2 Materials & Methods

5.2.1 Cell Culture

In preparation for the experiments, cells were revived from frozen cell banks into GSK proprietary growth media. Cells were incubated at 37°C, 5% CO₂, and 150 rpm in an orbital shaking incubator. Cells were sub-cultured every 3-4 days until sufficient volumes of cell culture were available for bioreactor inoculation at a suitable viable cell density. Each experiment contained cells from an identical number of passages to ensure the number of generations was kept constant and in line with experiments detailed in Section 4.2.1.

Four experiments (RS-E-MPC-N) were conducted in a 2L Single-use stirred-tank Univessel® SU (Sartorius) bioreactor with a 1.6L working volume to culture a GSK mAb-producing CHO cell line. Using a $1 \times 10^6 \text{ cell.mL}^{-1}$ seeding density, the process was operated under fed-batch conditions for 15 days with a feeding of a proprietary GSK feed on days 3, 6, 8, 10, and 13. Meanwhile, a supplementary feed of glucose (300 g.L⁻¹) was fed on days 3, 6, 7, 8, 9, 10, 13, and 14 with the volume of glucose fed determined by four different process control strategies in each experiment. During the first two experiments, a severe operating error created the opportunity to demonstrate the limits of process control. In detail, a cell culture dilution occurred on day 2 due to an unclosed valve. This caused an inadvertent addition of media, which was subsequently followed by the required removal of cell culture material to restore the working volume to 1.8L. This allowed for two process control scenarios to be tested: (1) reactive scheduling in the presence of a major operating error, RS-E, and (2) nonlinear model predictive control in the presence of a major operating error, MPC-E. Following this, two further experiments were run under otherwise normal operating conditions, albeit featuring two different process control scenarios: (3) open-loop optimisation in the absence of major operating errors, OLO-N, and (4) nonlinear mode predictive control in the absence of major operating errors, MPC-N. Details of (1)-(4) are outlined in following sections. PID controllers were used to maintain a fixed temperature of 35°C and an allowable pH range of 6.9-7.0 controlled using NaHCO₃ and CO₂ additions. Dissolved oxygen tension (DOT) was controlled to a setpoint of 50% using air and O₂ supplied via a sparger at a variable ratio starting at 20 ccm of air with no pure O₂, after which, O₂ is controlled depending on demand and was capped

at 300 ccm.. Antifoam (Gibco™ FoamAway™ Irradiated AOF, ThermoFisher Scientific) was supplemented to the cell culture and a constant agitation speed of 225 rpm was applied throughout the process.

5.2.2 Analytical Methods

In this research, only widely available process sensor technologies were utilised with the aim of showcasing how MPC can be implemented with existing technologies without retrofitting current manufacturing methods. More advanced analytical technologies may be (a) challenging to implement, (b) especially if they are offline, take longer to perform measurements when time-constrained decision-making is required during control, or (c) be difficult to replace if a process sensor becomes unavailable due to failure or miscalibration. In the latter case, process control can still provide inputs, but they may not be optimal due to deviations between the model and the process (Sommeregger et al., 2017). Cell culture samples were taken daily. A RAPIDPoint 500 Blood Gas Analyzer (Siemens Healthineers) was used to calibrate and validate the pH sensor using a single daily measurement. To measure culture viability, viable cell concentration, and total cell concentration, 0.5mL of the cell culture sample was mixed with equal parts of TrypLE™ Express enzyme (Thermo Fisher Scientific) to enable the dissociation of aggregated cells. This was then split into three samples of 0.17mL and incubated for 10 minutes at 37°C before all three samples were measured by a Vi-CELL XR Cell Viability Analyser (Beckman Coulter) and the results averaged. The remainder of the cell culture sample was centrifuged for 2 minutes at 800 x g. The final supernatant was extracted, and the pellet was discarded. From this, a Cedex Bio HT® Analyser (Roche) was used with three repeat samples of 0.1mL of the supernatant to obtain extracellular concentrations of glucose, lactate, glutamate, glutamine, ammonia, lactate dehydrogenase and product titre. Finally, an OsmoPro Micro-Osmometer (Advanced Instruments) was used to obtain the osmolality measurement using two repeat measurements. The remainder of the sample was stored at -80°C.

5.2.3 Mathematical Modelling

Robust and accurate mathematical models are a key tool towards applying process systems engineering tools. In this research, a mathematical model is used for process control of fed-

batch CHO cell culturing thus requiring (1) an operational strategy, and (2) implementation of this via a process control algorithm. During the first step, the model is utilised prior to process operation for *in silico* studies to perform a systems analysis. In the second step, the model is utilised during several model-based control strategies: using (1) open-loop optimisation, (2) reactive scheduling, or (3) model predictive control.

The success of the model-based control algorithm depends directly on several aspects including the accuracy and robustness of the process model. Further, for reactive scheduling and model predictive control, where the model is simulated and inputs are optimised during process operation, there are key considerations in the model choice. For example, linear data-driven models of fed-batch CHO cell cultures may be advantageous due to lower computational demands when solving optimisation problems online. Here the optimisation can be cast as a convex problem with clearer global minima (Luo, Kurian, et al., 2023). On the contrary, optimisation of non-linear, non-convex systems is more demanding requiring increased iterations whilst there is an increased risk of only finding local rather than global optimal solutions (Camacho et al., 2007). Despite this, control based on linear models may result in an unacceptable response including a large process model mismatch, oscillatory behaviour, or instability (Craven et al., 2014; Sendrescu, 2011; Tenny et al., 2004). In addition, due to the relatively slow process dynamics compared to chemical processes, and relative infrequency of control inputs, there is greater flexibility to use more computationally demanding process control algorithms. Finally, the risk of optimisations falling on local optima can be reduced with sound initial guesses and well-chosen constraints (Tenny et al., 2004). Previously, one of the greatest barriers to non-linear model predictive control of fed-batch CHO cell cultures was building a model that describes the process dynamics (Craven et al., 2014) but recent research has alleviated this (Kyriakopoulos et al., 2018; Tang et al., 2020) including that presented in Chapter 4.

Overall, mechanistic models have prior demonstrations for efficient evaluation, analysis, and optimisation of fed-batch CHO cell cultures (Kotidis, Demis, et al., 2019; Kotidis, Jedrzejewski, et al., 2019; Kyriakopoulos et al., 2018; Tang et al., 2020). Further, mechanistic models may be able to extrapolate to conditions that they have not been trained to (Fung Shek

et al., 2021). Therefore, such models are hypothesised to be suitable in the event of unexpected, process disturbances and may offer robust control.

Considering the above, in this chapter we utilise the mechanistic model designed for online application and presented in Chapter 4.

5.2.4 Process Control Implementation

As described in Section 5.2.1, three different process control methods are demonstrated in this research: open-loop optimisation (during the OLO-N experiment), reactive scheduling (during the RS-E experiment), and model predictive control (during the MPC-E and MPC-N experiments). These methods require optimisation to define the manipulated variables. Glucose can be highly influential in the cell's metabolism and thus was manipulated during closed-loop control in previous studies (Lu et al., 2013). Further, in Chapter 4, a model was developed showcasing the effects of manipulating the volume of glucose fed. Therefore, this research will only consider the volume of a feed with a fixed concentration of glucose as a manipulated variable. The controlled outputs could include process outputs such as product titre and will be defined in the following sections. Dynamic optimisation was performed in gPROMS ModelBuilder v7.1.1 (Siemens PSE) using the control vector parameterisation single-shooting (CVP_SS) solver for optimisation. For each process control algorithm in the remainder of this section, a dynamic optimisation problem (Eq. 25a) is solved across the finite time horizon, $T = [t_0, t_f]$ where t_0 is the initial time, and where t_f is the time horizon. For open-loop optimisation t_0 is equivalent to zero. For model predictive control and reactive scheduling, t_0 is the initial conditions and reset upon receipt of new data, and optimal control is re-performed.

$$\min_{u(t)} \phi(t_f) \text{ subject to} \quad \text{Eq. 25a}$$

$$\dot{x}(t) = f(x(t), y(t), u(t), v(t)) \quad \forall t \in T \quad \text{Eq. 25b}$$

$$x(t) = X_v(t), X_t(t), [mAb](t), [Glc](t), [Lac](t) \quad \text{Eq. 25c}$$

$$y(t) = mu_{growth}(t), mu_{death}(t), Q_{mAb}(t), Q_{Glc}(t), Q_{Lac}(t), Osm(t) \quad \text{Eq. 25d}$$

$$u(t) = F_{in,Glc}(t) \quad \text{Eq. 25e}$$

$$v(t) = F_{in}(t), F_{out}(t) \quad \text{Eq. 25f}$$

$$I(x(0), y(0), u(0), v(0)) \quad \text{Eq. 25g}$$

$$x_{min} \leq x(t) \leq x_{min} \quad \forall t \in T \quad \text{Eq. 25h}$$

$$y_{min} \leq y(t) \leq y_{min} \quad \forall t \in T \quad \text{Eq. 25i}$$

$$u_{min} \leq u(t) \leq u_{min} \quad \forall t \in T \quad \text{Eq. 25j}$$

Here, an objective function (Eq. 25a) containing the controlled variables is minimised using a process model (Eq. 25b) where $x(t)$ and $y(t)$ are the differential variables (Eq. 25c) and algebraic variables Eq. 25d) in the model respectively. Therein, X_v and X_t is the viable cell density and the total cell density, respectively. Whilst $[mAb]$, $[Glc]$, $[Lac]$ are the product titre and extracellular concentrations of glucose and lactate, respectively. Meanwhile, Eq. 25d lists the specific growth rate (u_{growth}), specific death rate (μ_{death}), specific productivity (Q_{mAb}), specific glucose utilisation rate (Q_{Glc}), specific lactate utilisation rate (Q_{Lac}), and extracellular osmolality (Osm). The only manipulated variable ($u(t)$) is the volume of glucose fed to the cell culture, $F_{in,Glc}$, as shown in Eq. 25e. Further, the fixed model inputs are the volume of feed in (F_{in}) and sample volume out (F_{out}) as displayed in Eq. 25f. The dynamic optimisation problem is also subject to initial conditions, I , for all process variables (Eq. 25g). Finally, constraints may be set on the differential variables (Eq. 25h), algebraic variables (Eq. 25i), or manipulated variables (Eq. 25j).

5.2.4.1 Open-loop Optimisation

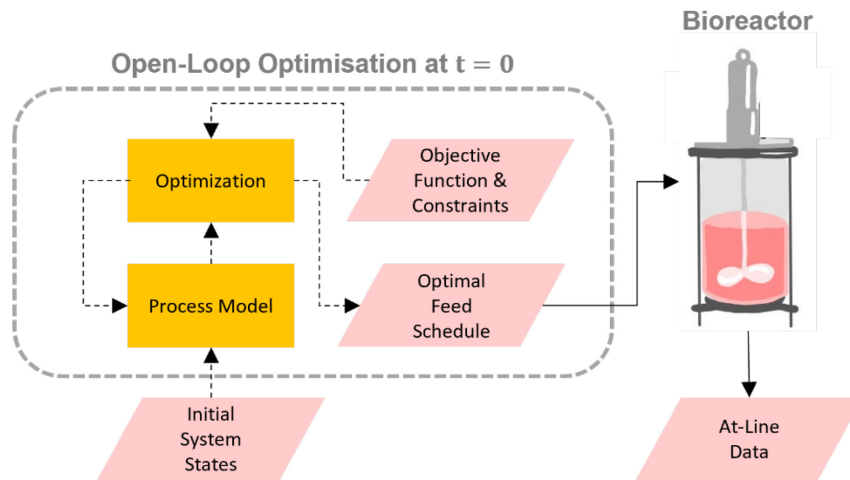


Figure 27 Open-loop optimisation architecture used for the OLO-N experiment. OLO-N: open-loop optimisation experiment in the absence of major operating errors.

The most basic to implement of the process control techniques in this study is open-loop optimisation shown in Figure 27. Initial system states required by the model are defined. Model-based dynamic optimisation is then performed prior to operation or at $t=0$ using an objective function, constraints, and initial guesses for the optimum manipulated variables. A warm start was used for the optimisation whereby the initial guesses are equal to the best-performing feeding strategy from previous experiments (e.g., experiment MF-L, see Chapter 4). This is particularly useful when dealing with non-convex, highly non-linear systems where optimisations can frequently get caught at local optima. From this, an optimal feed schedule is defined which remains fixed throughout the process. During open-loop optimisation, there is a clear lack of utilisation of at-line data. Despite successful *in silico* (Kappatou et al., 2020) and experimental (Kotidis, Jedrzejewski, et al., 2019) demonstrations of open-loop-optimisation, it is expected that relative to other proposed control techniques, open-loop optimisation may be more sensitive to significant process model-mismatch, or process disturbances. Here, the lack of feedback may lead to a misinformed feeding schedule, thus yielding sub-optimal cell culture conditions.

5.2.4.2 Reactive Scheduling

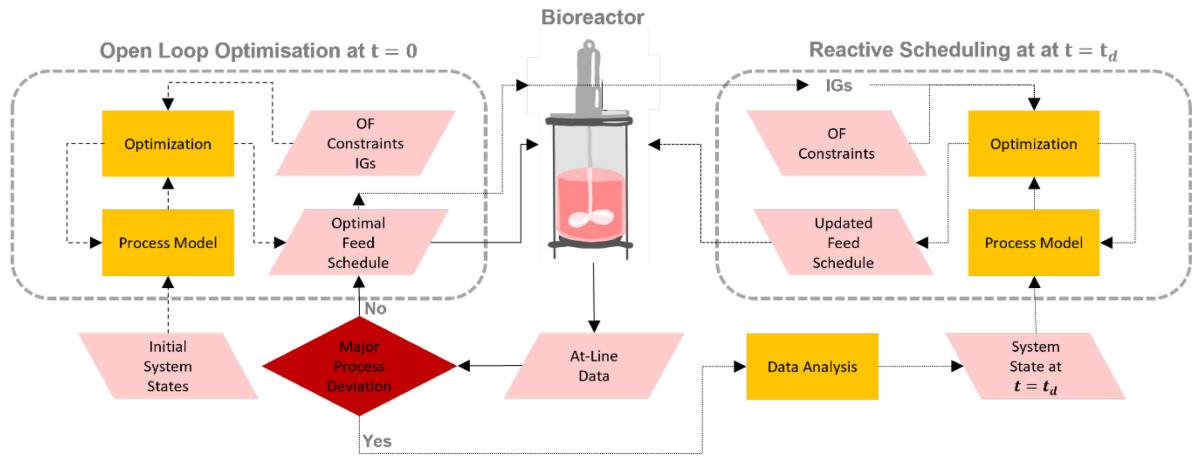


Figure 28 Reactive scheduling architecture used for the RS-E experiment. OF: Objective function. IGs: Initial guesses. t_d : time of major process deviation. RS-E: reactive scheduling in the presence of a major operating error.

To increase process robustness, reactive scheduling can be used in the event of unforeseen disturbances or deviations to recover the process towards optimality using an updated operation schedule (Figure 28). Implementation of reactive scheduling begins with a pre-defined schedule, in this case derived using open-loop optimisation at time = 0. Upon accounting for a major process deviation, the process can be corrected by updating the schedule e.g., using receding horizon-based optimisation (Kopanos & Pistikopoulos, 2014). Unlike model predictive control, which is applied continuously (and therefore more computationally expensive) irrespective of the level of process deviation, reactive scheduling acts only after a diagnosis of a process deviation to minimise the effect of this deviation on the remaining process schedule (Ruiz et al., 2000). Here, a deviation in process conditions was caused by an operational error leading to harsh changes in process inputs described in Section 5.2.1. As such model-based dynamic optimisation was performed at the time of the disturbance (t_d) using the most recent measurements, with initial guesses taken from the open-loop optimisation optimal feed schedule, and a new feed schedule implemented which accounts for the disturbance. Although previously proposed and tested *in silico* for control of glycosylation in fed-batch CHO cell culturing (Kotidis et al., 2021), reactive scheduling has not yet been experimentally demonstrated in open research during fed-batch production of mAbs using CHO cell culturing. Rather most focus has been on implementing model predictive control.

5.2.4.3 Model Predictive Control

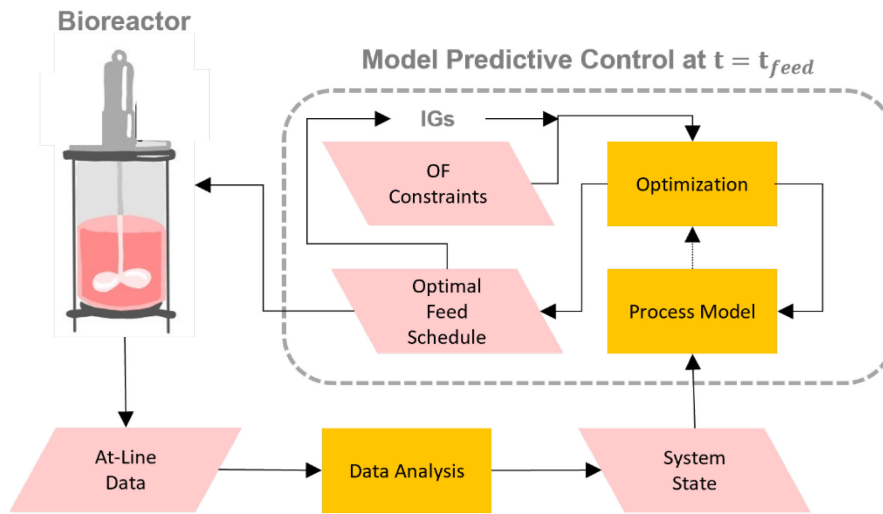


Figure 29 Model predictive control architecture used for experiments MPC-E and MPC-N. OF: Objective function. IGs: Initial guesses. MPC-E: model predictive control in the presence of a major operating error.

MPC-N: model predictive control in the absence of major operating errors.

Model predictive control utilises a receding horizon whereby upon receipt of new process measurements, at a feeding time, the model's initial states are reinitialised, and model-based dynamic optimisation is performed to minimise an objective function under a set of constraints, based on the most recent process state. At the first iteration, the initial guesses are equal to the best-performing feeding strategy from previous experiments (MF-H, see Chapter 4). For subsequent iterations, the initial guesses are taken from the previous iteration of optimisation. In this study, as the length of the cell culture is fixed, the controller will utilise a receding horizon.

5.2.5 Operational Strategy Development

Two key steps were used to develop a suitable control strategy namely understanding the system dynamics using sensitivity analysis and performing *in silico* optimisation. Global Systems Analysis within gPROMS ModelBuilder v7.1.1 (Siemens PSE) was used to perform sensitivity analysis and distinguish the relative importance of each of the inputs used in the model in terms of their influence on the model outputs. This variance-based sensitivity analysis method is only suggested where the number of input factors is under 20 (Saltelli et al., 2008).

In this case, experiments were performed in an early-stage process development environment with limited opportunities for feeding glucose, only on days 3, 6, 7, 8, 9,10, 13, and 14. As such, the total number of input factors is equivalent to the total number of feeding opportunities, here equal to 8. Further constraints were set on the system so that the volume of glucose fed did not exceed the maximum volume of glucose fed during experiments used to tune the model.

As described in Section 4.2.4, sensitivity analysis is used to determine first order and total effect indices for each parameter normalised by the variance in response (Saltelli et al., 2010). The first order effect for each parameter is attributed to the contribution to the uncertainty for that parameter alone and can be used for factor prioritisation (Saltelli & Tarantola, 2002), whereas the total effect is the combination of linear and nonlinear contributions from all parameters on the model's uncertainty and can be used to facilitate factor fixing (Saltelli & Tarantola, 2002). Quasi-random Sobol sampling is a sampling method that yields no overlapping points (Sobol, 1976). It was used to generate 4096 sets of feeding strategies to be used during sensitivity analysis with bounds set dependent on the confidence intervals obtained during parameterisation. The first order and total effect indices are estimates of the Sobol indices that converge to their true value as the number of samples increases (Saltelli et al., 2010). As such, the number of parameter sets must be sufficiently large such that the estimated Sobol indices remain stable.

One of the key advantages of model-based process development is the opportunity to reduce experimental effort. Here, following sensitivity analysis, *in silico* optimisation is used to predict how different variations in the objective function formulation affect the process output. Following Eq. 25, dynamic optimisation was performed in gPROMS ModelBuilder v7.1.1 (Siemens PSE) using the control vector parameterisation single-shooting (CVP_SS) mathematical solver.

5.3 Operational Strategy Design

Designing an operational strategy for model-based process control will help formulate an optimal objective function and constraints. In other words, which variables will be controlled in the objective function by the manipulated variables. As mentioned, this is not trivial, especially for a highly non-linear system. To minimise experimental efforts, there are several computational tools used within this workflow (e.g, global sensitivity analysis and model-based optimisation) to test the operation strategy *in silico* before experimental implementation. During analysis, the system was constrained loosely initially to minimise deviations away from the conditions which the model was trained upon thus minimising extrapolation into operating regions where the system behaviours are unknown. To do so, a constraint on the manipulated variable is imposed as shown in Eq. 26 prior to sensitivity analysis.

$$F_{in,Glc}^{min} \leq F_{in,Glc} \leq F_{in,Glc}^{max} \quad \text{Eq. 26}$$

Here, $F_{in,Glc}^{max}$ was set to the maximum feed rate of glucose that occurred during the feeding schedule performed within the model-building experiments (MF-N, MF-L, and MF-H presented in Figure 16 and Chapter 4). Further, $F_{in,Glc}^{min}$ was set to zero rather than the minimum feed rate of glucose that occurred during the feeding schedule performed within the model-building experiments outlined in Section 4.2.1. This will allow any control strategy to act appropriately and, in the event of glucose over-feeding, no glucose may be fed to prevent exacerbating the negative effects caused by excessive extracellular glucose concentration and resulting hyperosmolality. Both bounds chosen for Eq. 26 mean that there may be instances during analysis where the model is extrapolating outside the region for which it was trained. However, setting wide bounds increases the ability to explore the full design space. Further, for a non-linear system, it will be particularly difficult to ascertain whether a prediction is outside the model interpolation region (Masana & Baranyi, 2000). Setting bounds within the minimum and maximum feed rate of glucose at each time point may still induce extrapolation due to the non-convexity of non-linear systems (Fisher et al., 2022). Given this, extrapolation can be unavoidable and is incurred frequently during the deployment of non-linear models (Dong & Ma, 2022; Shen & Meinshausen, 2023). However, the use of a mechanistic model for systems analysis may allow for a greater degree of extrapolation (Fisher et al., 2022).

5.3.1 Input-Output Analysis

Using global systems analysis, the system behaviour can be predicted when the feeding schedule is changed as shown in Figure 30 to Figure 34. Each point represents the output of each modelled state on the sample point directly after every feeding event. As glucose is fed on days 3, 6, 7, 8, 9, 10, 13, and 14, it is of interest to decipher how the system may react at the next sample point directly prior to the next feed on days 6, 7, 8, 9, 10, 13, 14, and 15.

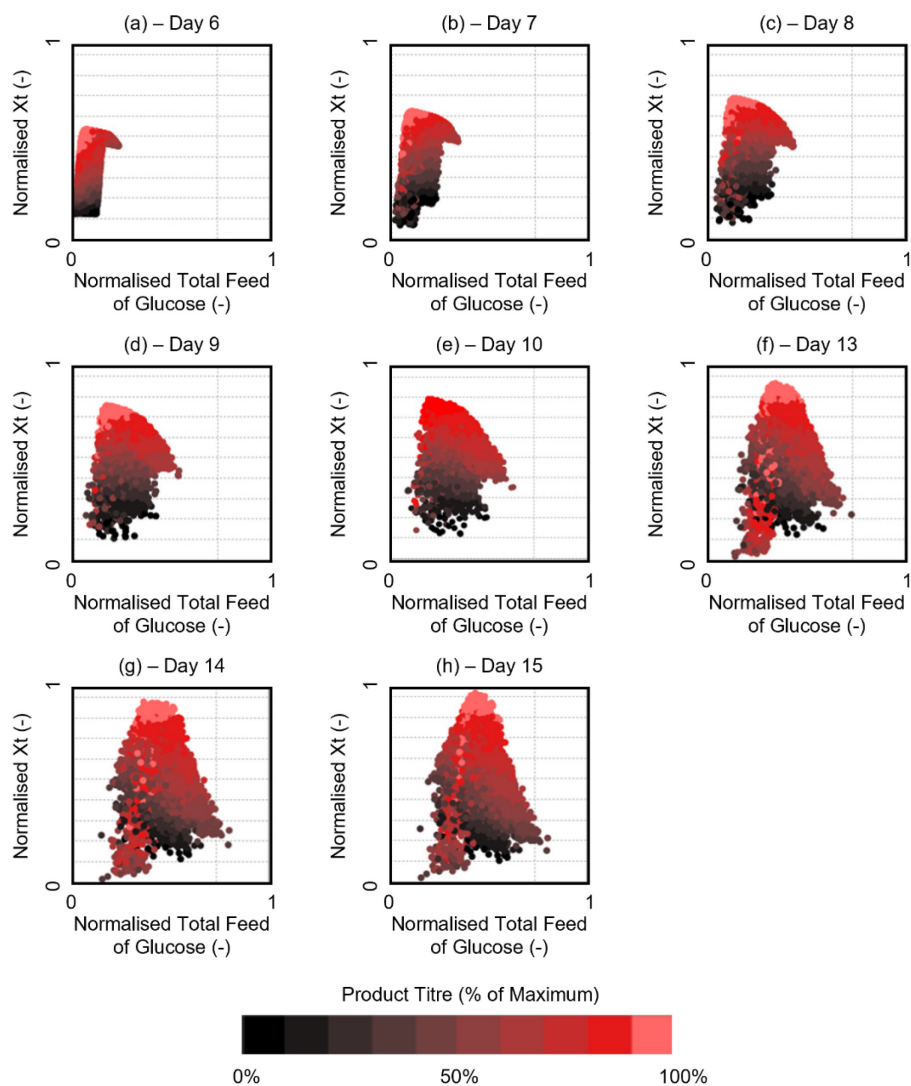


Figure 30 Input-output analysis comparing, at each day, the total feed volume of glucose to the total cell density and product titre. Normalised using min-max normalisation.

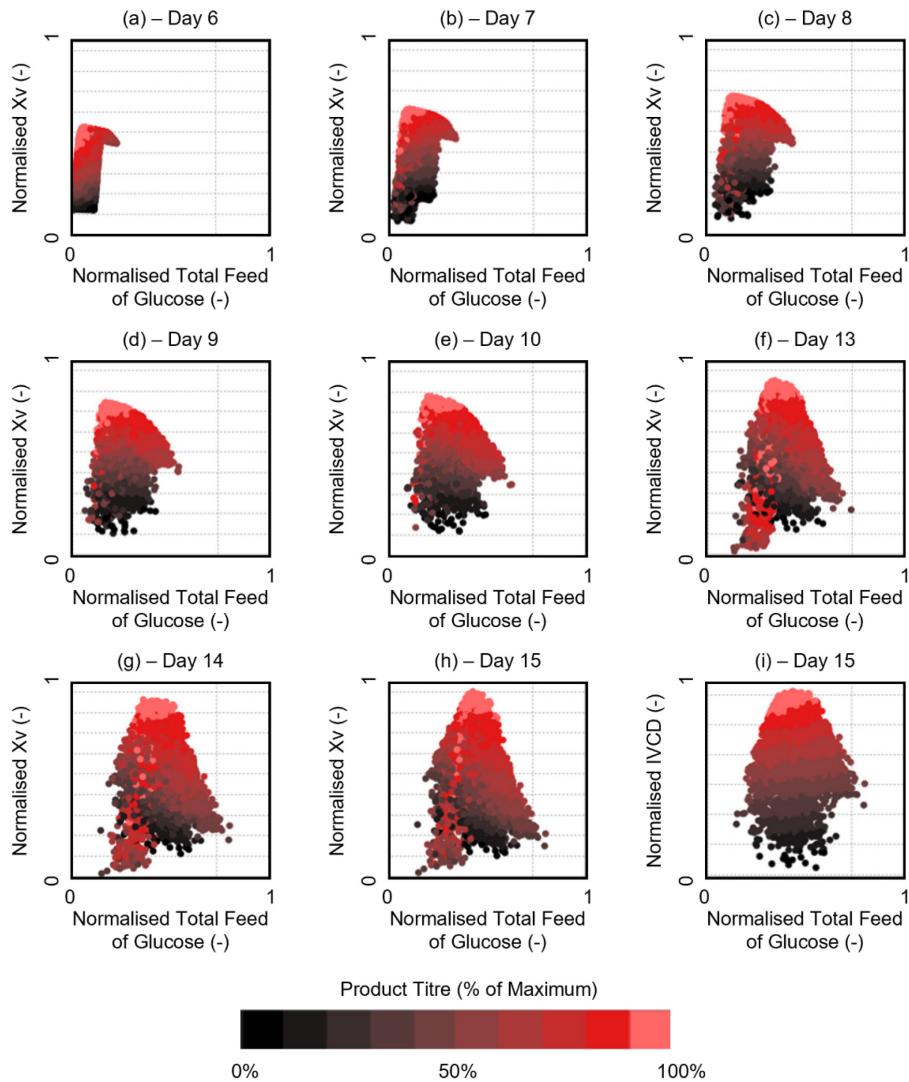


Figure 31 Input-output analysis comparing, at each day, the total feed volume of glucose to the viable cell density and product titre. Normalised using min-max normalisation.

The behaviour of total cell density and viable cell density is highly similar as can be compared in Figure 30 and Figure 31. As the days progress, the width of the characterisation space across the x-axis increases where the range in the total feed has increased and multiple different feed schedules achieve similar cell concentrations. There is also, quite a large width of the characterisation space across the y-axis. This could be explained by the wide constraints on the volume of glucose fed, with the minimum volume being zero to allow correction in the event of over-feeding. Therefore, many of these in-silico simulations will be extrapolating outside of the conditions they were parameterised upon. In practice, the effects of glucose starvation may

vary during production relative to the model predictions presented here. On day 10, it is clearly observable that the higher the viable cell density, the greater the product titre. This relationship is obvious when observing the integral viable cell density (Figure 30i) and intuitive as the specific production rate of mAb in the process model is approximated by a single parameter, m_{mAb} . On Day 13 onwards, high product titres are achieved at lower values of cell density. However, this does not indicate a weak performing process. Instead, this is due to the modelled effects of cell death rate, and growth rate reducing from day 13 onwards. It also corroborates previous studies wherein maximising integral viable cell density is a useful objective for a model-based optimisation strategy (Koumpouras & Kontoravdi, 2012) rather than solely viable cell density. Therein, maximising IVCD independently did not lead to maximum [mAb] and vice versa albeit endpoint values were within 10% of each other thus considered under experimental error. As such IVCD and [mAb] will both be considered for use within the objective function.

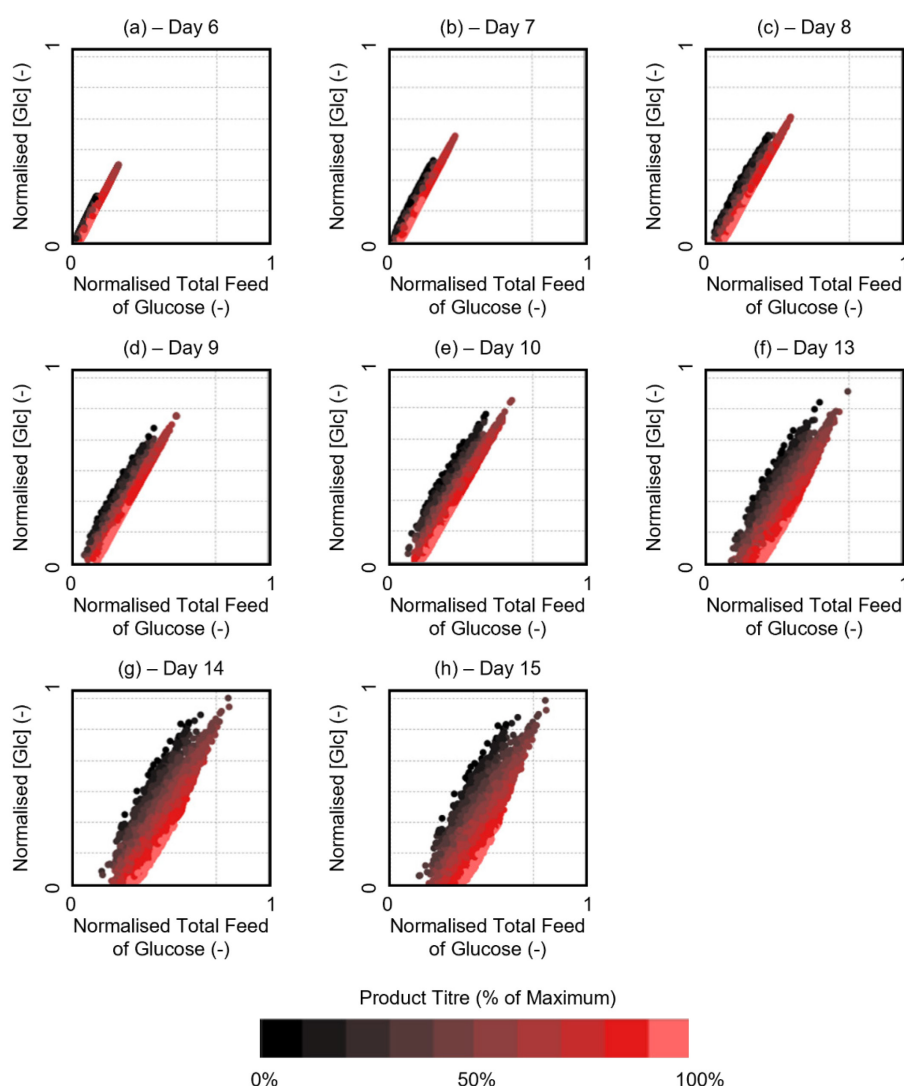


Figure 32 Input-output analysis comparing, at each day, the total feed volume of glucose to the extracellular concentration of glucose and product titre. Normalised using min-max normalisation.

In Figure 32, the positive correlation between increased total feed volume of glucose and extracellular glucose concentration is clear and intuitive. However, this result also shows that at higher concentrations of glucose, the product titre is lower. In all instances, only at lower extracellular concentrations of glucose is the product titre at the upper 10% of the maximum highlighting the importance of preventing over-feeding. This effect was observed during previous experiments (MF-N, MF-L, and MF-H) and built into the model whereby increased concentration of glucose leads to a higher osmolality, thus lower cell growth. Lower cell growth may lead to this effect worsening as the concentration of viable cells decreases leading to a greater excess of unconsumed glucose. However, at lower concentrations of glucose, cell

growth will be limited negatively affecting product titre. This is indicated in Figure 32 where the left-hand side of the characterisation space always achieves lower product titre and vice versa on the right. This highlights the importance of tight control of glucose which itself is not trivial (Craven et al., 2014; Domján et al., 2020). Minimisation of the manipulated variables is common within model predictive control where achieving control with minimal control action is desirable to lead to smooth guidance (Schwenzer et al., 2021). To achieve this effect in this study, and to prevent overfeeding, the volume of glucose fed into the process will be minimised.

Lactate is widely known to be an inhibitory metabolite and as such previous studies have attempted to suppress lactate production using closed-loop control (Matthews et al., 2016) or media supplementation (Buchsteiner et al., 2018). One of the key challenges is that, despite inhibitory effects, the production of lactate is associated with rapid cell growth during the exponential growth phase (Torres et al., 2018). This is highlighted in Figure 33 which shows the relationship between the total feed volume of glucose, extracellular lactate concentration, and product titre. In particular, Figure 33a highlights how increased lactate concentration is predicted when higher product titres are achieved. This could be because increased lactate secretion is typically associated with increased cell growth, and therefore higher product titre (Torres et al., 2018). However, a valuable operating strategy may include inducing successive consumption of lactate to prevent growth inhibition following production (Torres et al., 2018). The challenge here would usually be defining and predicting when lactate consumption occurs. However, the model developed in (Chapter 4) predicts lactate consumption and production. This can be seen in particular in Figure 33 where the high product titres initially are present at increased concentrations of lactate in Figure 33c and subsequently occur at lower concentrations in Figure 33f - Figure 33h due to the inbuilt lactate consumption kinetics within the mechanistic model. This presents a useful opportunity to utilise model-based control of lactate. However, solely minimising extracellular lactate concentration or the peak lactate concentration runs the risk that growth may be indirectly reduced (Torres et al., 2018). To lessen this risk, a reduction to prolonged exposure to heightened amounts of lactate is proposed.

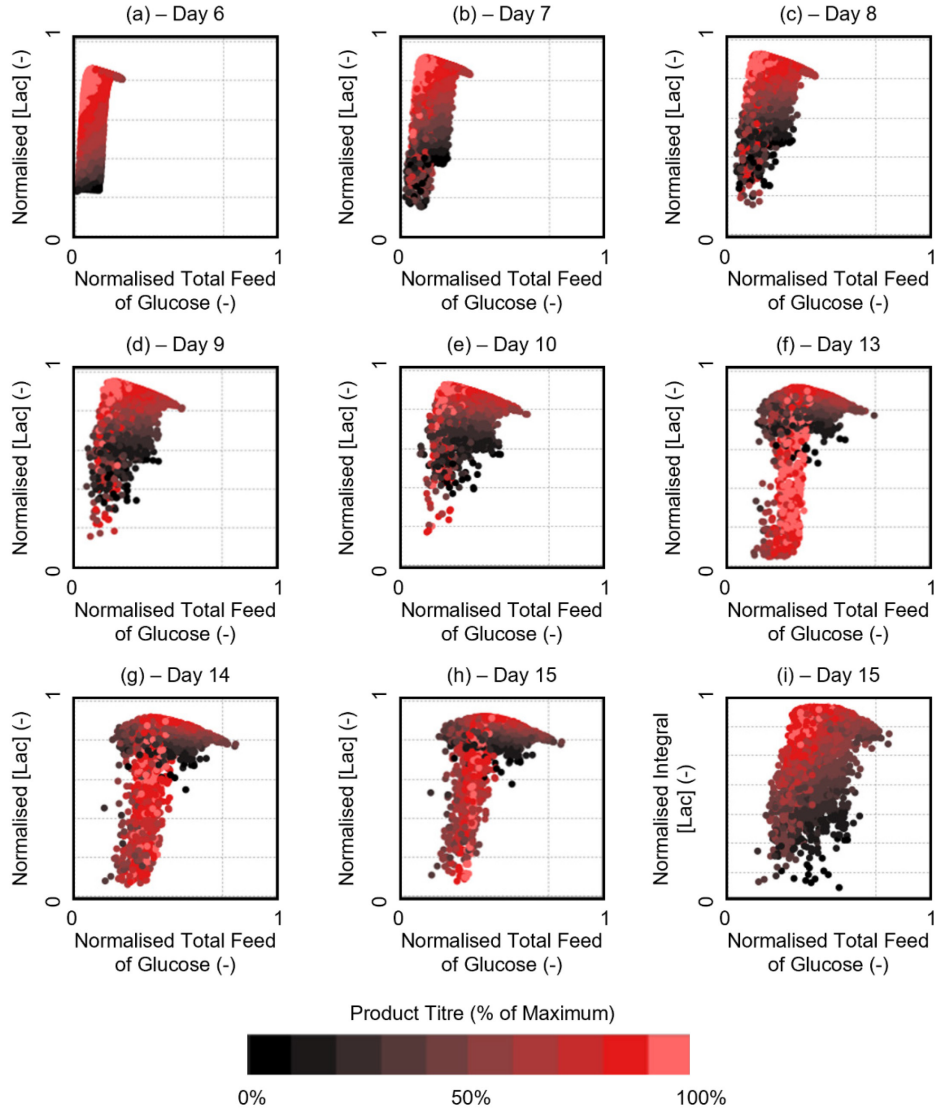


Figure 33 Input-output analysis comparing, at each day, the total feed volume of glucose to the extracellular concentration of lactate and product titre. Normalised using min-max normalisation.

To quantify the exposure to extracellular lactate concentration across the cell culture period, the integral of this was calculated. Here it is postulated that minimising integral lactate concentration may allow for a high peak lactate concentration to occur, thus high growth, and subsequently force lactate consumption if possible, preventing prolonged exposure to high lactate concentrations. Here, the trapezium rule was used to estimate the integral extracellular lactate concentration ($IntLac(t)$, Eq. 27) from experimental data during each feeding interval, starting at $(t - 1)$ and ending at (t) . Further, the variation in integral lactate is shown in Figure

33i which shows only a gentle decrease is required in integral lactate to avoid negatively impacting product titre. This can be applied via a suitably weighted objective function.

$$IntLac(t) = IntLac(t - 1) + \frac{([Lac](t) + [lac](t - 1))\Delta t}{2} \quad \text{Eq. 27}$$

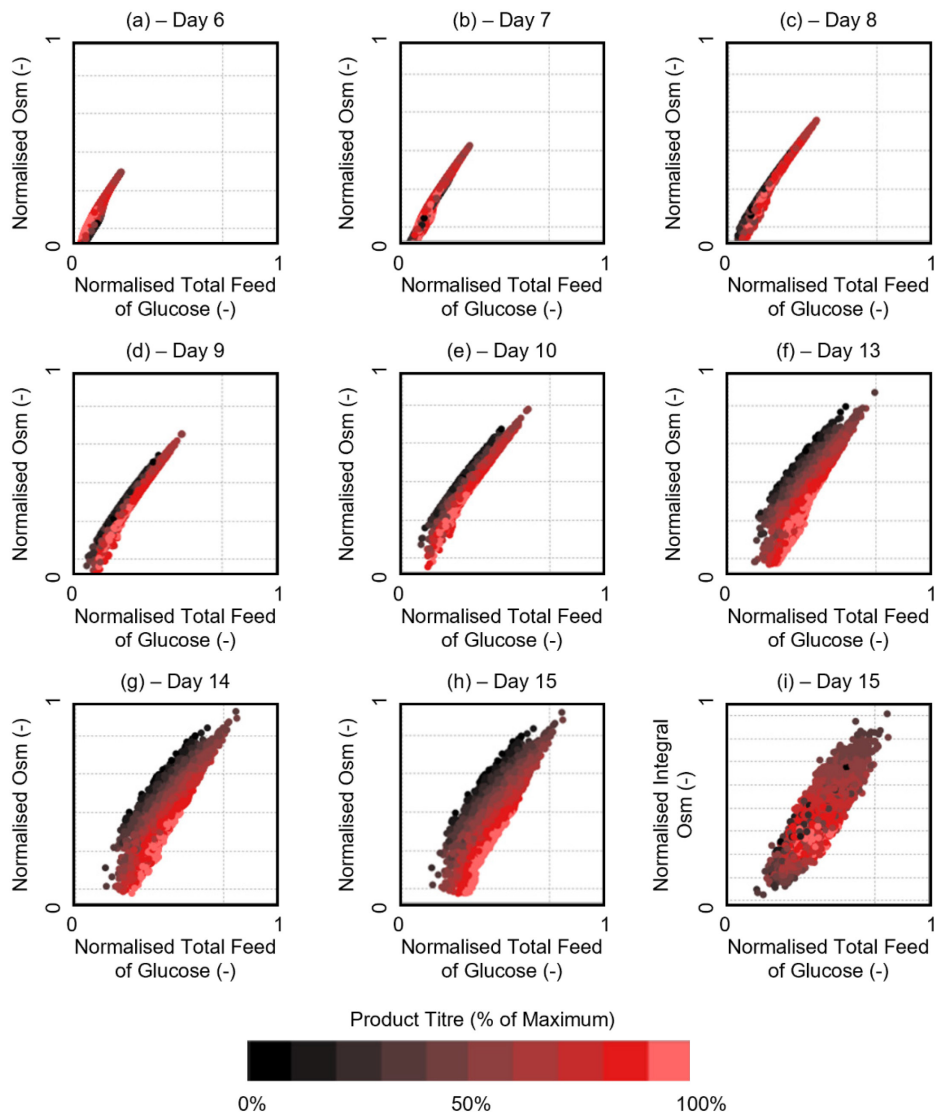


Figure 34 Input-output analysis comparing, at each day, the total feed volume of glucose to the extracellular osmolality and product titre. Normalised using min-max normalisation.

Finally, Figure 34 shows the relationship between osmolality, total feed volume of glucose and product titre. Due to the osmolality being a function of extracellular concentration, the results and explanations reflect Figure 32 closely. In detail, the higher the total feed volume of glucose, the higher the osmolality. As observed in previous research (Alhuthali et al., 2021), high osmolality negatively affects growth, hence lower predicted product titres are observed at higher osmolality. In addition, during the model-building, it was observed experimentally, that integral osmolality (Eq. 28) was correlated to the overall growth rate (Figure 23). Therefore, this will also be minimised during model-based optimisation. Similarly, to the integral lactate, this will have to be weighted appropriately so the minimisation is not so forceful that it negatively impacts product titre (Figure 34i).

$$IntOsm(t) = IntOsm(t - 1) + \frac{(Osm(t) + Osm(t - 1))\Delta t}{2} \quad \text{Eq. 28}$$

5.3.2 Input-Output Sensitivity Analysis

To further analyse the effects of variations in the feeding schedule on the system behaviour, sensitivity analysis was used to calculate (a) the first order sensitivity indices, (b) the total order sensitivity indices, and (c) the ratio between the two to check what proportion the factor has interactions with the other factors. During this analysis, each day of glucose feed was treated as a separate factor affecting each of the six modelled states, and the sensitivity indices were taken at the next time of sampling. In detail, as glucose is fed on Days 3, 6, 7, 8, 9, 10, 13, and 14, then it is of interest to decipher how the system has reacted at the next sample point directly prior to the next feed on Days 6, 7, 8, 9, 10, 13, 14, and 15. Thus, for the latter set of days, the sensitivity indices are displayed for each state in Figure 35 to Figure 37. The only constraint applied was on the feed rate of glucose whilst Sobol sampling (a quasi-random sampling method) was used, assuming a normal distribution, to produce a set of feeding schedules such that no overlapping points were generated (Sobol, 2001). A sensitivity threshold was set at 0.05 to filter parameters that have a negligible contribution to model variability (Kotidis, Demis, et al., 2019; Saltelli et al., 2004; X. Y. Zhang et al., 2015). This choice is discussed further in

Section 4.2.4. Figure 35 shows how the first order effects change throughout the cell culture duration. It is immediately apparent, that the volume of glucose fed on day 3 is one of the most influential inputs on the success of the process (Figure 35a-Figure 35h). Further, the volume of glucose fed early in the culture has the greatest effect on the final product titre (Figure 35h). This indicates the priority in deploying predictive control to ensure that the feeds of glucose early on consider simulated process trajectories. Meanwhile, altering these trajectories later in the process using only glucose will be challenging as indicated by the low S_i on all days apart from day 3. This is expected as fed-batch mammalian-based process dynamics typically exhibit large time delays.

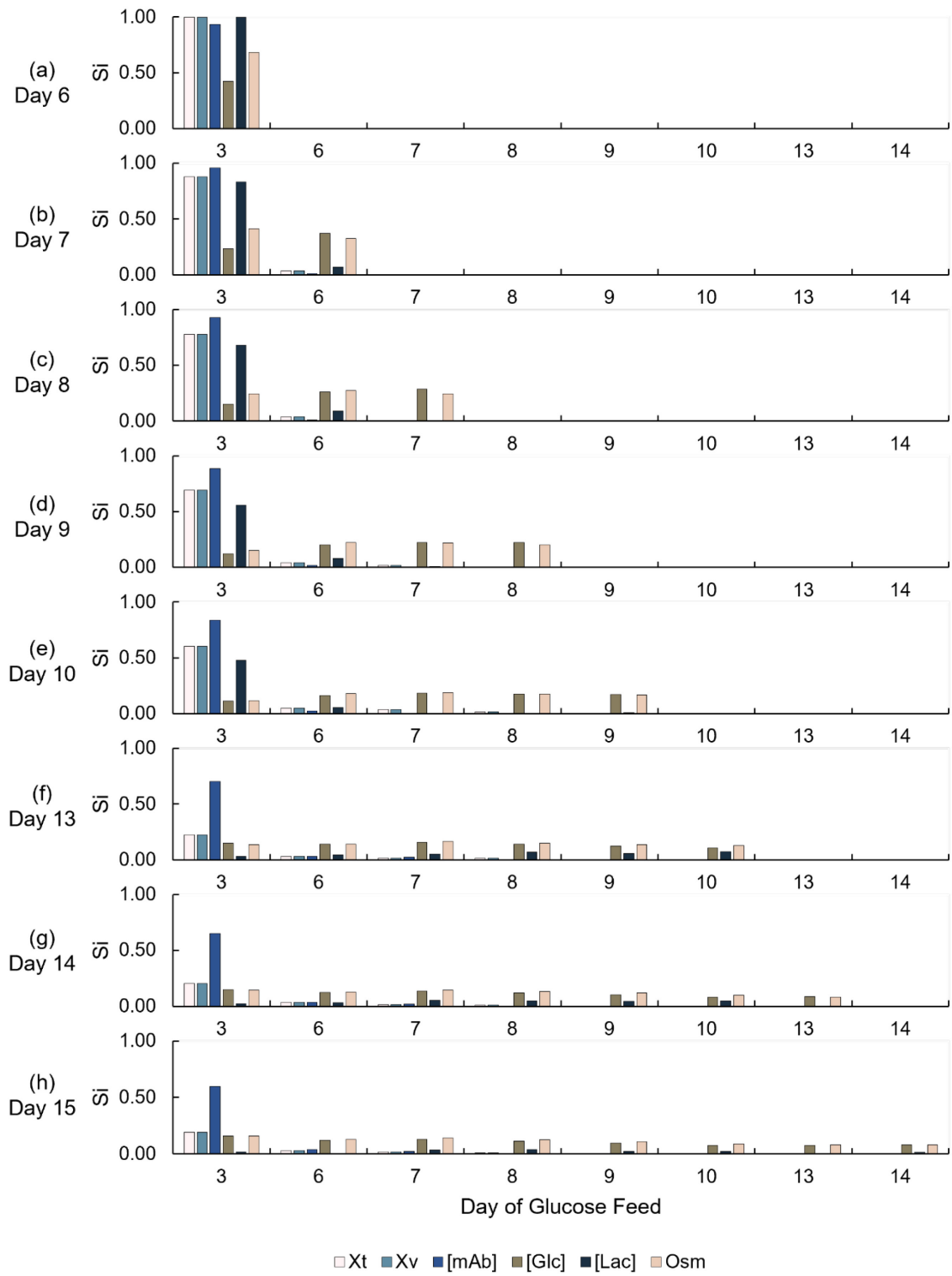


Figure 35 First order sensitivity indices showcasing the effects of each scheduled glucose feed on the process outputs at the sample time after feeding, and prior to the subsequent feed. Normalised using min-max normalisation.

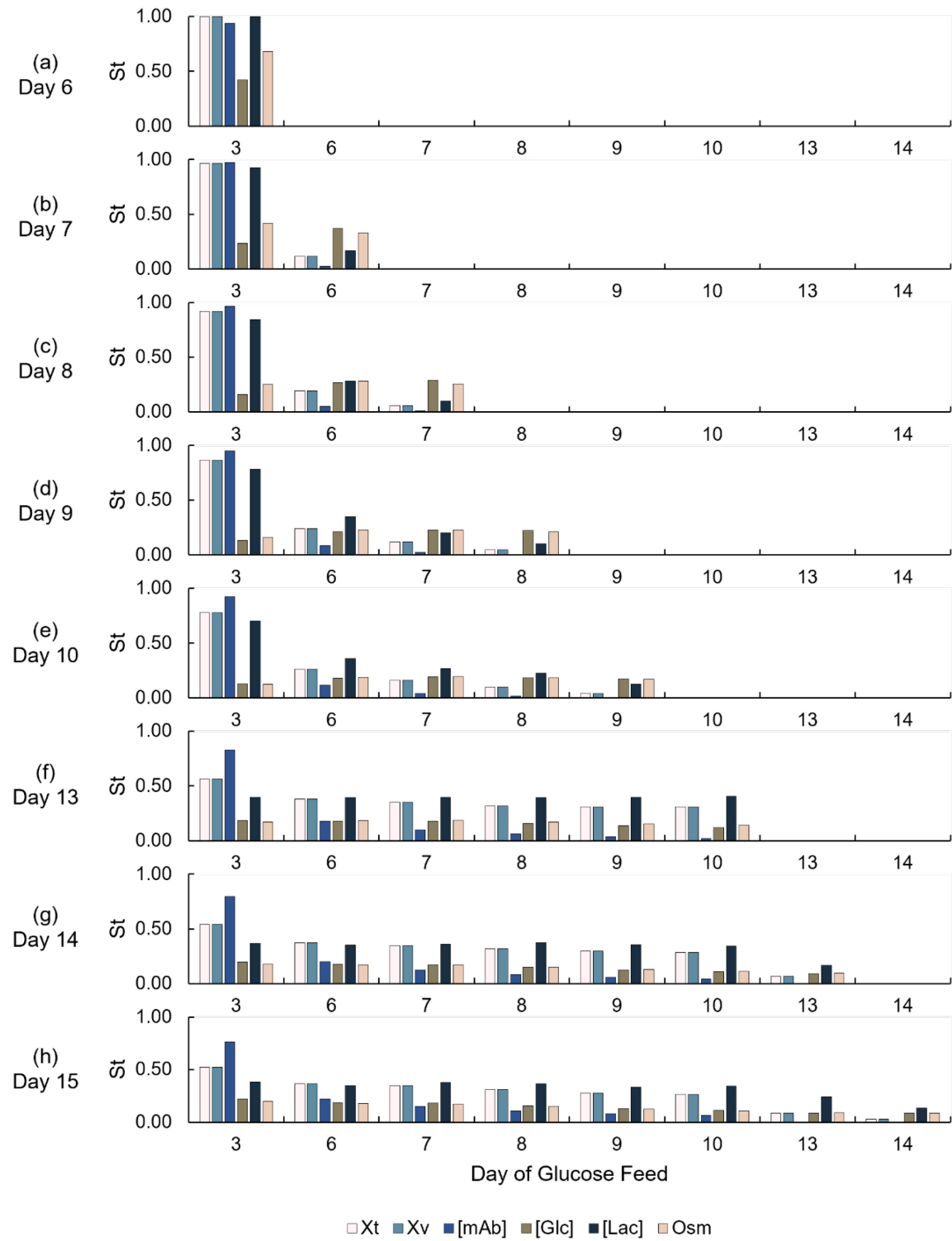


Figure 36 Total effects indices showcasing the effects of each scheduled glucose feed on the process outputs at the sample time after feeding, and prior to the subsequent feed. Normalised using min-max normalisation.

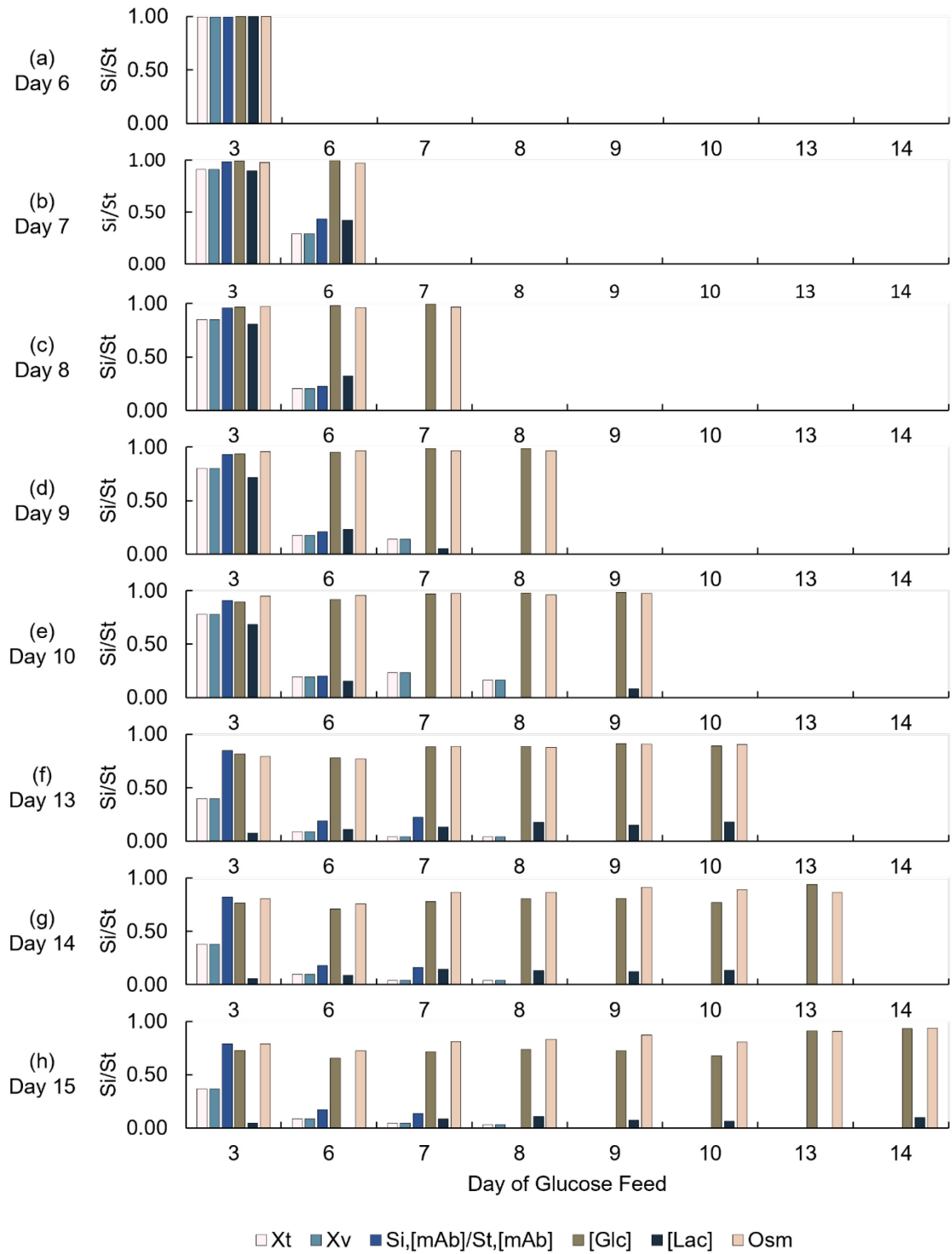


Figure 37 Ratio between the first order sensitivity indices and total effects indices showcasing the effects of each scheduled glucose feed, and the value of the process outputs at the sample time after feeding, and prior to the next feed. Normalised using min-max normalisation.

Figure 36 shows the total order sensitivity indices of the feed rate of glucose on each day and can be used to showcase the input's interactions with outputs. Here, the feed rate of glucose continues to affect total cell density and viable cell density throughout the process to varying degrees. This is expected as glucose is a primary carbon source of cells and previous input-output analysis (Figure 32). That said, it should be noted that glucose is also the only growth-limiting metabolite fed that the model considers. There may be other metabolites fed that are excluded from the model and this analysis that are biologically growth-limiting. Looking across the process, the feed rate of glucose influences to a much lesser amount the product titre, up until the feed of glucose on day 10. After day 10, the final endpoint product titre does not appear to be significantly influenced by the feed of glucose interactions (Figure 36h). This also displays the slow process dynamics present in the process between the feed of glucose, cells growing, and expressing product. Further, this highlights the use case of using a predictive control method where on day 10, the controller is accounting for future predictions of product titre on day 15. Interestingly, the effects of glucose have a steady minor effect on lactate throughout the process across all days. Biologically lactate is inhibitory to growth and has been controlled in the past with indirect effects on viable cell density and product titre (Matthews et al., 2016; Van Beylen et al., 2020). Further, as the model utilised in the present research shows how lactate may be influenced by glucose feed volume, then extracellular lactate concentration could be included within the control objective. Finally, Figure 37 presents the ratio between first order and total-order effects. A value close to one indicates no significant second order interactions due to the feed of glucose on the specified output feed (X. Y. Zhang et al., 2015). Here it is evident that throughout the process both glucose and osmolality are primarily affected by first order effects of glucose feed. This is an expected result as the volume of glucose fed will directly affect the extracellular concentration of glucose which instantaneously increases the osmolality.

Overall, the sensitivity analysis highlights (a) the importance of correct feeding during early feed points, (b) the advantage of predictive control methodology which can act based on future projections, (c) the potential to include variables that biologically affect product titre within the objective function but perhaps show little influence by glucose within the current model. On the latter point, it should be highlighted that the model approximates the system, as such it may describe some process variables better than others. If, for example, it describes lactate

accurately but not product titre, then control of lactate through optimisation may lead to control of product titre. As such, having multiple controlled variables may increase the likelihood of effective control.

5.3.3 *In silico* Optimisation

In parallel with the above analysis, several *in silico* optimisations were performed within gPROMS (Siemens PSE) varying objective functions detailed in Table 16. In detail, dynamic optimisation was sought to manipulate the volume of glucose feed to optimise various operational strategies (S1-S6) containing variables discussed in Section 5.3.2 and included maximising monoclonal antibody titre (mg/L), maximising monoclonal antibody amount (mAb_{mg}, mg), maximising integral viable cell density (cell/L), minimising integral osmolality (mM. kg H₂O), minimising integral extracellular lactate concentration (mM), or maximising a multi-variable objective function.

Table 16 *In silico* optimisation strategies. +: Maximise. -: Minimise.

	[mAb]	mAb _{mg}	IVCD	IntOsm	IntLac	IntFin
S1	+					
S2		+				
S3			+			
S4				-		
S5					-	
S6	+	+	+	-	-	-

The dynamic optimization strategies were performed using the same time horizon at which the model was developed within (360 hours). Minimal constraints were applied including on the feed rate of glucose to minimise and prevent excessive control action whilst maintaining conditions like that which the model was parametrised within, the osmolality to prevent hyper-osmolality (Alhuthali et al., 2021), and the minimum extracellular concentration of glucose to ensure the model did not deviate from conditions for which it was parameterised. The results

of the optimisations can be seen in Figure 38 which shows the normalised endpoint values of key variables.

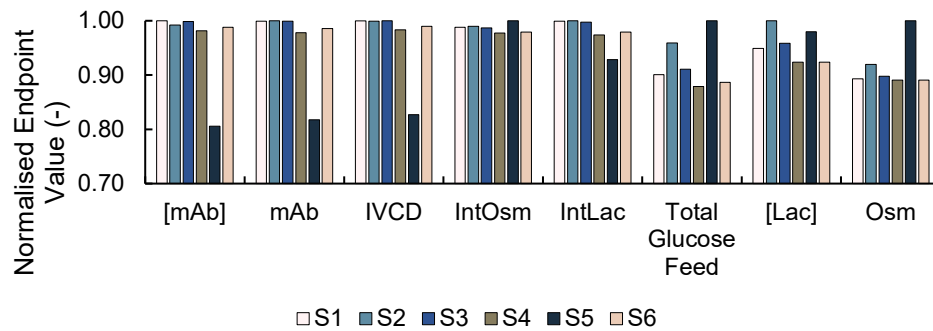


Figure 38 Results from *in silico* optimisation. Normalised using min-max normalisation based on the max of each endpoint value.

In Figure 38, there is very little difference between S1 and S2 in [mAb], mAb_{mg}, IVCD, IntOsm, and IntLac with deviations below the typical measurement error of each of these variables. However, as S1 is maximising a concentration, the total glucose feed is also lower preventing cell culture dilution. Consequently, this causes the endpoint [Lac] and Osm to be lower for S1. For S3, maximising IVCD performs very similarly to maximising [mAb] and this indifference between S3 and S1 corroborates with similar research (Koumpouras & Kontoravdi, 2012). When minimising solely IntOsm in S4, this leads to a marginal decrease in IntOsm relative to other strategies, and a slightly lower [mAb], mAb_{mg}, IVCD versus S1 and S2. The worst-performing strategy was S5, minimising IntLac. Here, as lactate is a by-product of cell metabolism, minimising IntLac, has caused cell growth to be minimised thus S5 achieves lower IVCD and therefore a lower [mAb]. S6 has one of the lowest total amounts of glucose fed and achieves one of the lowest endpoint values for [Lac] and Osm. In addition, during S6, some of the lowest values of IntLac and Int Osm are achieved versus other strategies. This may promote favourable cell culture conditions due to low growth inhibition. Meanwhile, for S6, IVCD and [mAb] were close to the best-performing strategies. Overall, S6 provides a well-rounded strategy aiming to promote high cell growth (maximising IVCD), high product titre (maximising [mAb]), and favourable cell culture conditions (minimising IntLac, IntOsm,

and IntFin). As such this strategy was taken forward and formalised in Eq. 29 awith *in silico* trajectories of the cell culture performing optimally shown in Figure 39. In Eq. 29a, each term is scaled (using α_i) based on the order of magnitude of the endpoint values of each controlled variable in the objective function. As was the case here, often the weights of non-linear model predictive controllers are chosen using trial and error (Vallerio et al., 2014). This strategy ensured that IVCD was prioritised whilst the remaining controlled variables still were influential on the objective function value. Further development could focus on more exact tuning methods (Vallerio et al., 2014).

$$\min_{u(t)} \left(-\alpha_1 [mAb] - \alpha_2 \int_{t_0}^{t_f} X_v + \alpha_3 \int_{t_0}^{t_f} Osm + \alpha_4 \int_{t_0}^{t_f} [Lac] + \alpha_5 \int_{t_0}^{t_f} F_{in,Glc} \right) \text{ s.t.} \quad \text{Eq. 29 a}$$

$$\dot{x}(t) = f(x(t), y(t), u(t), v(t)) \quad \forall t \in T \quad \text{Eq. 29b}$$

$$x(t) = X_v(t), X_t(t), [mAb](t), [Glc](t), [Lac](t) \quad \text{Eq. 29d}$$

$$y(t) = \mu_{growth}(t), \mu_{death}(t), Q_{mAb}(t), Q_{Glc}(t), Q_{Lac}(t), Osm(t) \quad \text{Eq. 29e}$$

$$u(t) = F_{in,Glc}(t) \quad \text{Eq. 29f}$$

$$v(t) = F_{in}(t), F_{out}(t) \quad \text{Eq. 29g}$$

$$I(x(0), y(0), u(0), v(0)) \quad \text{Eq. 29h}$$

$$[Glc]_{min} \leq [Glc]; Osm(t) \leq 450 \quad \forall t \in T \quad \text{Eq. 29i}$$

$$0 \leq F_{in}(t) \leq F_{in,Glc}^{max} \quad \forall t \in T \quad \text{Eq. 29j}$$

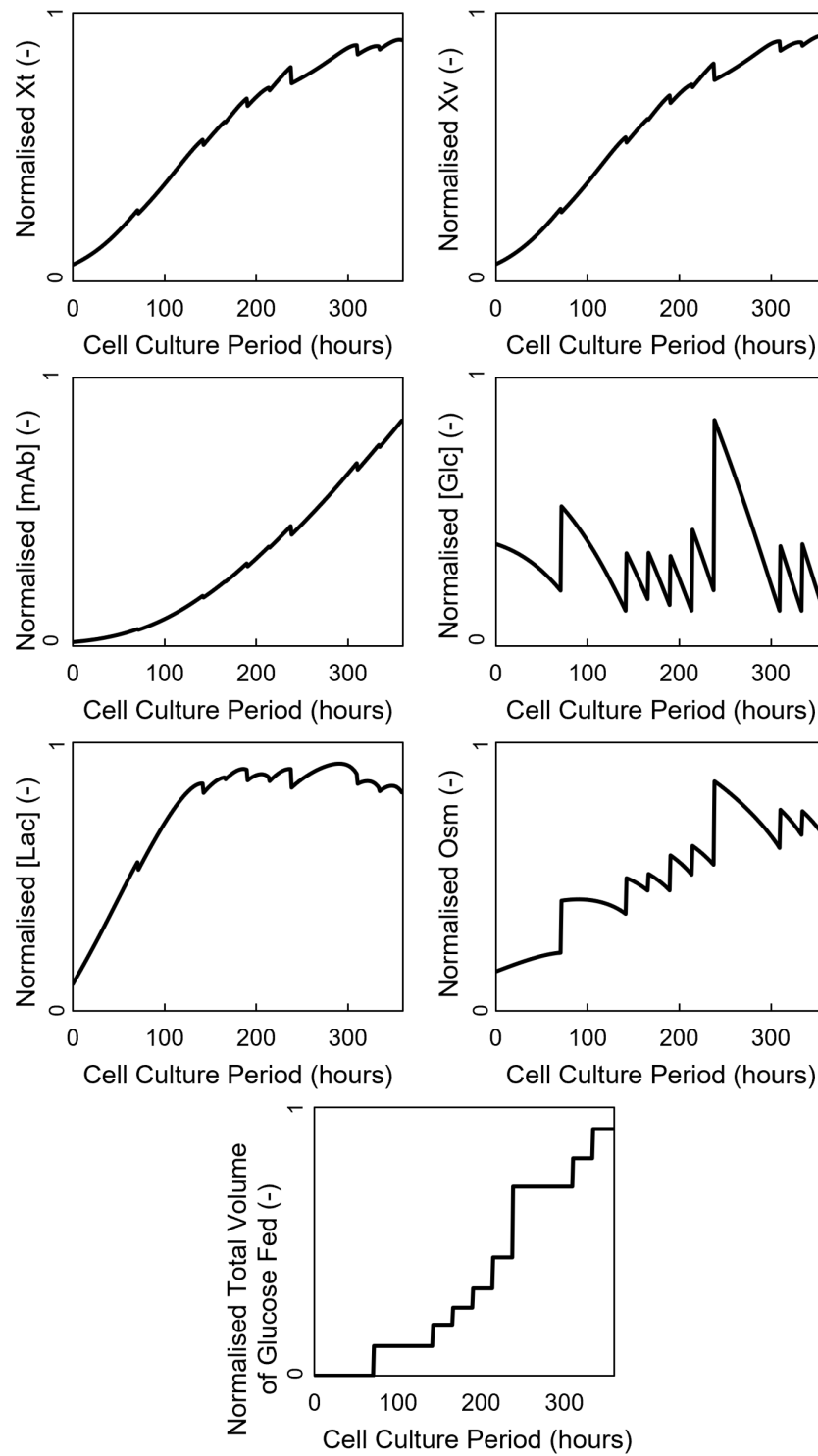


Figure 39 Optimal feeding schedule and *in silico* output trajectories for S6. Normalised using min-max normalisation based on a maximum value on the y-axis.

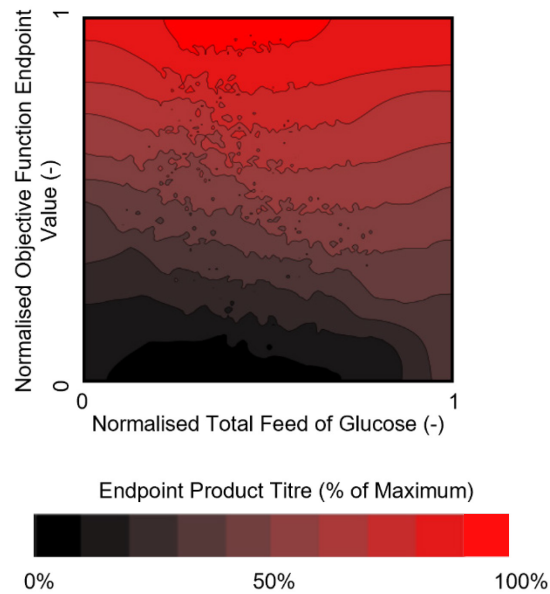


Figure 40 Relationship between manipulated variables, objective function, and product titre. Normalised using min-max normalisation.

In addition, the relationship between the total volume of glucose fed, the objective function and the endpoint product titre can be seen in Figure 40. Here there is a clear region of optimality where the endpoint product titre is highest. However, the non-linearity of the optimisation is also clear in parts where the relationship is possibly non-convex. As such, this may pose difficulty during optimisation where local optima may be reached. To improve the chance of successfully converging to a global optimum, warm starts are used during optimisations to provide initial guesses of the manipulated variables based on previous optimisations (Logist et al., 2010). In addition, very few hard constraints are used to help ensure the feasibility of the optimisation during closed-loop control (Schwenzer et al., 2021).

5.4 Results

5.4.1 Variation in Manipulated Variable

The variation in the amount of glucose fed each day depending on the feeding strategy, including compared to the model-building experiments, is shown in Figure 41. In all cases, the concentration of glucose was fixed, and the feed of all other components was kept fixed.

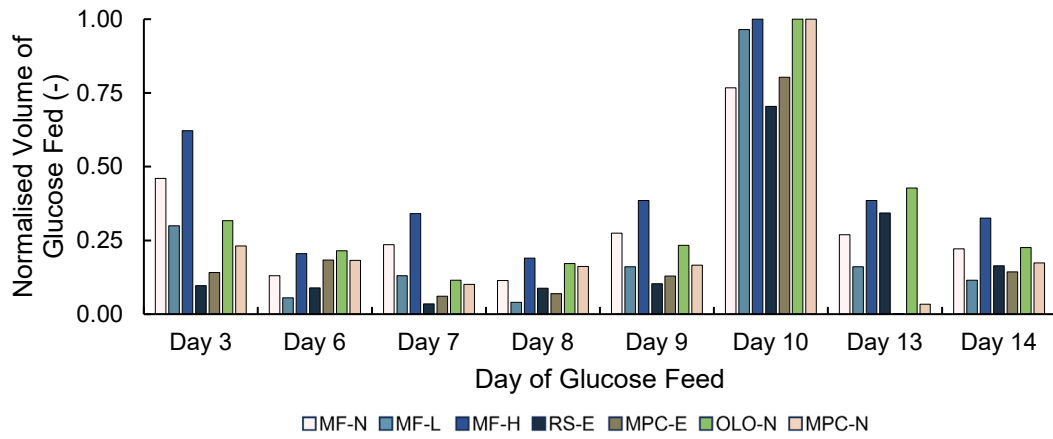


Figure 41 Variation in daily glucose feed due to the process control strategy. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose. RS-E: reactive scheduling in the presence of a major operating error. MPC-E: model predictive control in the presence of a major operating error. OLO-N: open-loop optimisation experiment in the absence of major operating errors. MPC-N: model predictive control in the absence of major operating errors. Normalised using min-max normalisation based on the overall maximum volume of glucose fed.

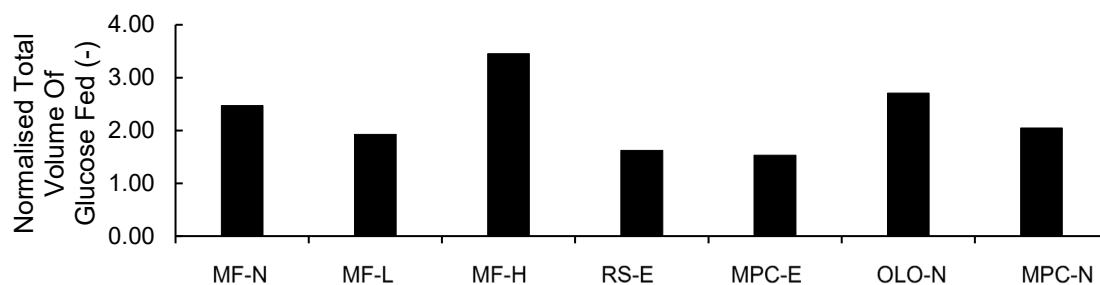


Figure 42 Variation in total glucose feed due to the process control strategy. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a high amount of glucose. RS-E: reactive scheduling in the presence of a major operating error. MPC-E: model predictive control in the presence of a major operating error. OLO-N: open-loop optimisation experiment in the absence of major operating errors. MPC-N: model predictive control in the absence of major operating errors. Normalised using min-max normalisation in Figure 41 then totalled in this Figure.

In Figure 41 it is visible that Day 10 has the highest volume of glucose fed. This is due to the highest cell density occurring during this time and the next possible feed being 3 days away. To further evaluate the differences in the volume of glucose fed between process control strategies, the mean and standard deviation were calculated. This was based on the volume of glucose fed each day, as well as on the total volume of glucose fed throughout the course of the cell culture. The results of this analysis are shown in Table 17.

Table 17 Variance and standard deviation of the volume of glucose fed across process control strategies.

	μ	σ
Day 3	0.309	0.183
Day 6	0.152	0.061
Day 7	0.145	0.107
Day 8	0.119	0.057
Day 9	0.207	0.098
Day 10	0.891	0.129
Day 13	0.231	0.170
Day 14	0.196	0.070
Total Glucose Feed	2.252	0.678

Table 17 reinforces that on Day 10 the largest volume of glucose is fed on average whilst the largest standard deviation occurs on Day 3 and Day 13. On Day 3, differences in feed can be mainly put down to the relatively large volume of glucose fed in the high (MF-H) glucose

feeding strategy (Figure 41). On Day 13, a large standard deviation occurs due to the low glucose feed in the model predictive control strategies (MPC-E and MPC-N) as visible in Figure 41. Specifically, the closed-loop model predictive control utilised in MPC-E (MPC with human error) and MPC-N (MPC under normal operation) feed close to zero as is visible in Figure 41. This may be caused by over-feeding on day 10 highlighting the drawback of the models' predictive capability during late-stage cell culture. Improved performance is expected with shorter feeding intervals.

5.4.2 Reactive Scheduling vs NLMPC During an Operational Error

The trajectories comparing model predictions versus experimental observations of RS-E (reactive scheduling) and MPC-E (non-linear model predictive control) are shown in Figure 43. During both experiments, an operational error was incurred as described in Section 5.2.1. In MPC-E, the model predictions are updated prior to each decision by changing the state to the experimentally observed value in a moving horizon manner. However, in RS-E, the model is only updated once immediately after the operational error occurs. It was anticipated that more frequent updates would give a more accurate predictive capability. However, as can be seen in Figure 43 and in more detail in Figure 44, the absolute percentage errors of RS-E and MPC-E are comparable. This is perhaps because the operational error created a situation where the model is extrapolating such that the advantage of model-predictive control is not as significant. Alternatively, this may be a drawback of minimising the number of metabolites modelled. For example, glucose may not be the only growth-limiting factor here and is, in any case, in abundance here.

In Figure 44, there is an apparent difference between RS-E and MPC-E most notably in the model predictions for glucose where RS-E appears to predict poorly with large absolute percentage errors. This is because in RS-E the control strategy does not include a state reset at every decision interval, only when the operational error occurs. Therefore, the plant model mismatch is not corrected.

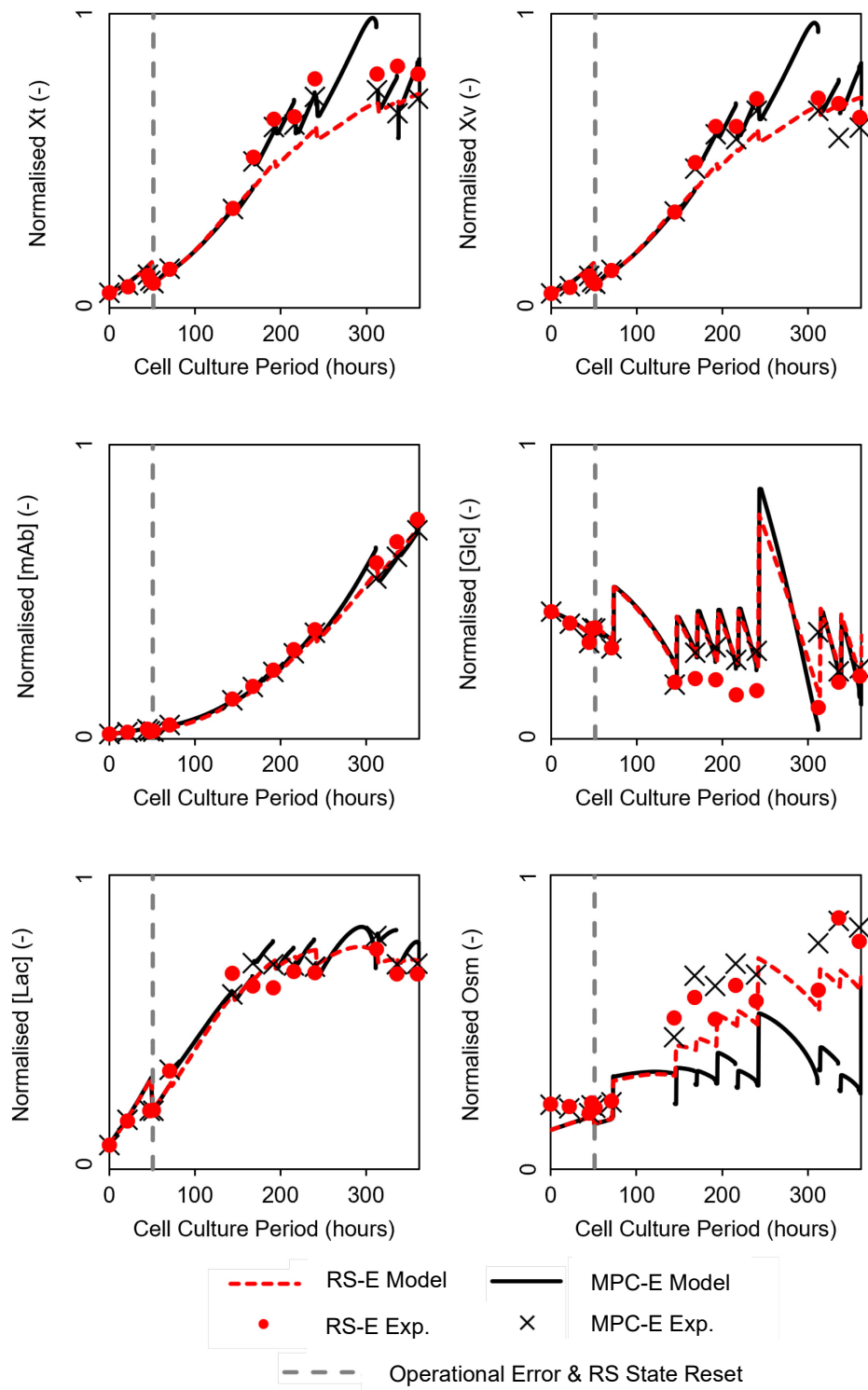


Figure 43 Trajectories of model predictions versus experimental observations when deploying reactive scheduling (RS-E) and non-linear model predictive control (MPC-E) during an operational error. Normalised using min-max normalisation.

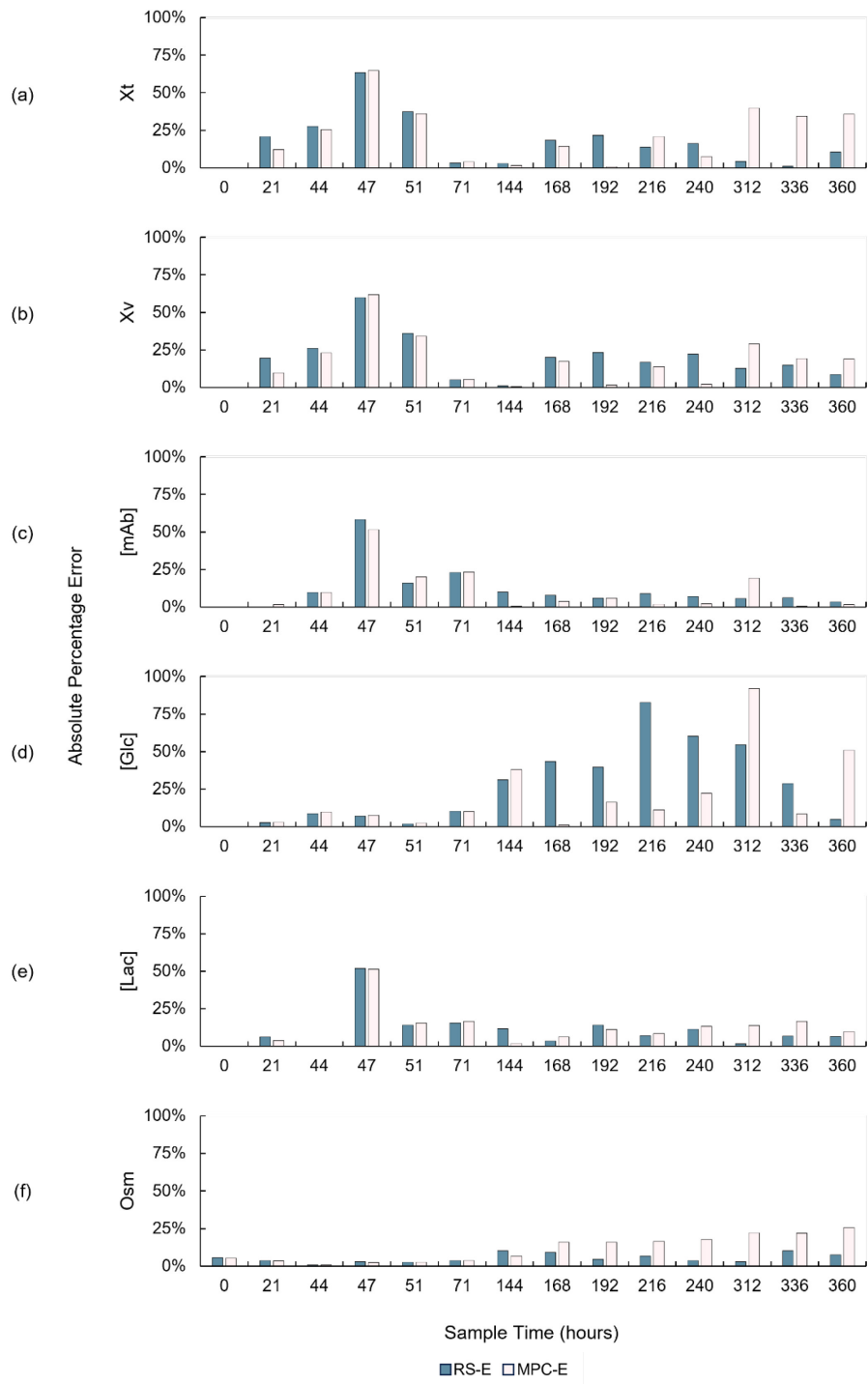


Figure 44 Absolute percentage error between model predictions versus experimental observations when deploying reactive scheduling (RS-E) and non-linear model predictive control (MPC-E) during an operational error.

Table 18 shows the R^2 which shows the fraction of model variability is explained by the experimental data. In addition, the RMSE represents the error between the model and the experiment. Finally, the PCC shows the correlation between the model and the experiment. The plant model mismatch present in MPC-E as highlighted in Table 18 where the R^2 is particularly low for osmolality in MPC-E. In addition, the R^2 for [Glc] is also particularly low in MPC-E. Here, this error is exacerbated by the difference between the measurement of [Glc] and the predicted value at 312 hours on Day 13 shown in Figure 43. Interestingly, the model predicts [mAb] well indicated by the high R^2 and PCC values. The poor predictive capability during MPC-E could indicate why the process control strategy fails to show a benefit over reactive scheduling as the latter requires less effort experimentally to obtain measurements and computationally to optimise after the operational error rather than at every decision point.

Table 18 Model performance when deploying reactive scheduling (RS-E) and non-linear model predictive control (MPC-E) during an operational error. RMSE: root mean square error in units of measurements for Xt (cells.mL⁻¹), Xv (cells.mL⁻¹), [mAb] (mg.L⁻¹), [Glc] (mM), [Lac] (mM), Osm (mOsmol.Kg⁻¹). PCC: Pearson correlation coefficient.

	RS-E			MPC-E		
	R^2	RMSE	PCC	R^2	RMSE	PCC
Xt	0.98	1.77E+06	0.99	0.96	1.69E+06	0.98
Xv	0.96	1.28E+06	0.98	0.94	2.30E+06	0.97
[mAb]	1.00	61.68	1.00	0.99	86.95	0.99
[Glc]	0.86	5.47	0.93	0.25	9.05	0.50
[Lac]	0.94	6.37	0.97	0.96	5.81	0.98
Osm	0.95	24.08	0.97	0.75	60.31	0.87

5.4.3 Open-loop Optimisation vs NLMPC during Normal Operation

The trajectories comparing model predictions versus experimental observations of OLO-N (open-loop optimisation) and MPC-N (non-linear model predictive control) are shown in Figure 45. These experiments were performed under normal operating conditions as described in Section 5.2.1. Further, in OLO-N, the model predictions are continuous, and the initial states are based on the experimentally observed conditions when the cell culture period is equal to zero. In MPC-N, the model predictions are updated prior to each decision by changing the state to the experimentally observed value. The absolute percentage errors between the model predictions and the experimentally observed results are shown in Figure 46.

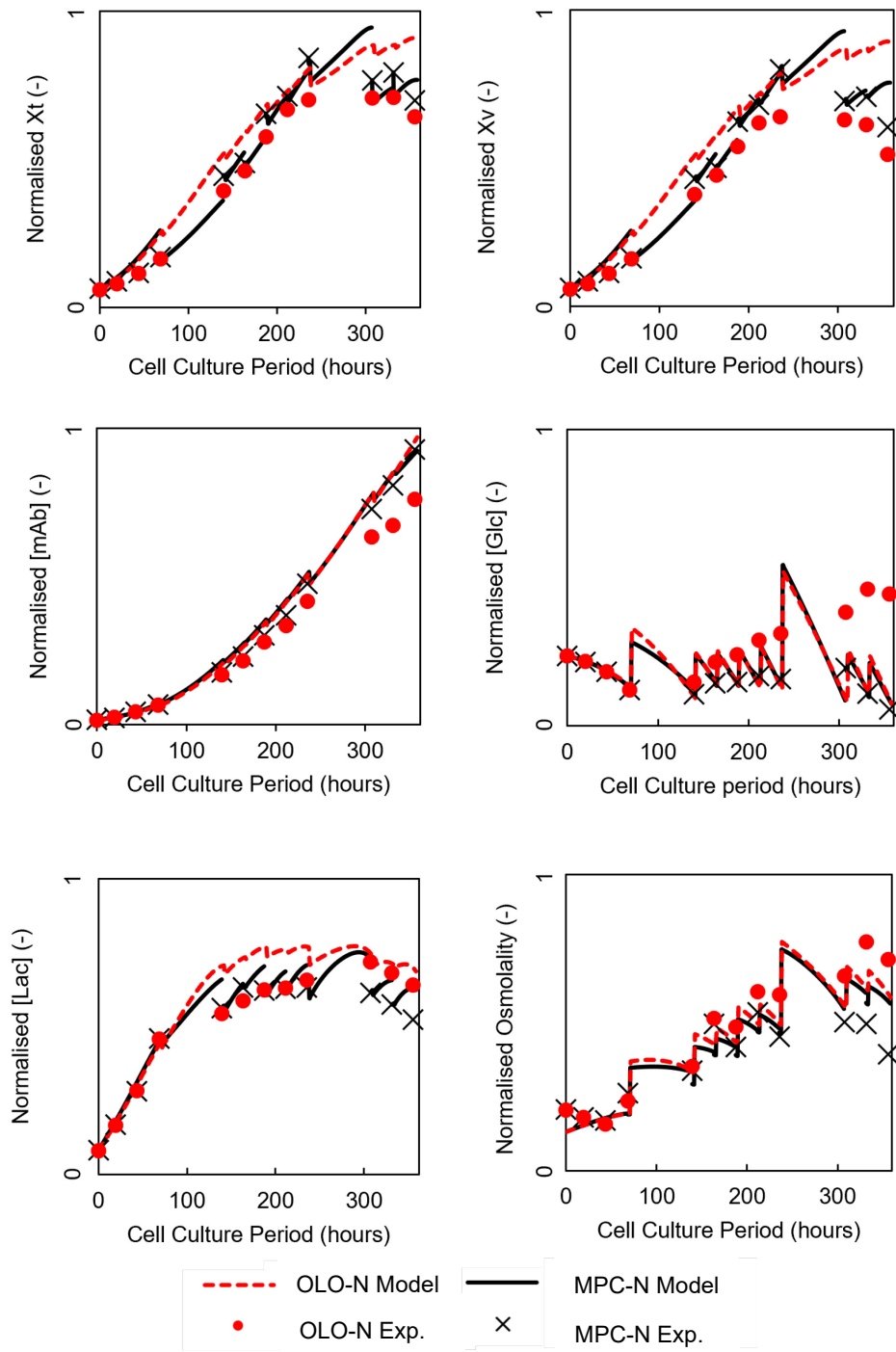


Figure 45 Trajectories of model predictions versus experimental observations when deploying open-loop optimisation (OLO-N) versus non-linear model predictive control (MPC-N) during normal operating conditions. Normalised using min-max normalisation.

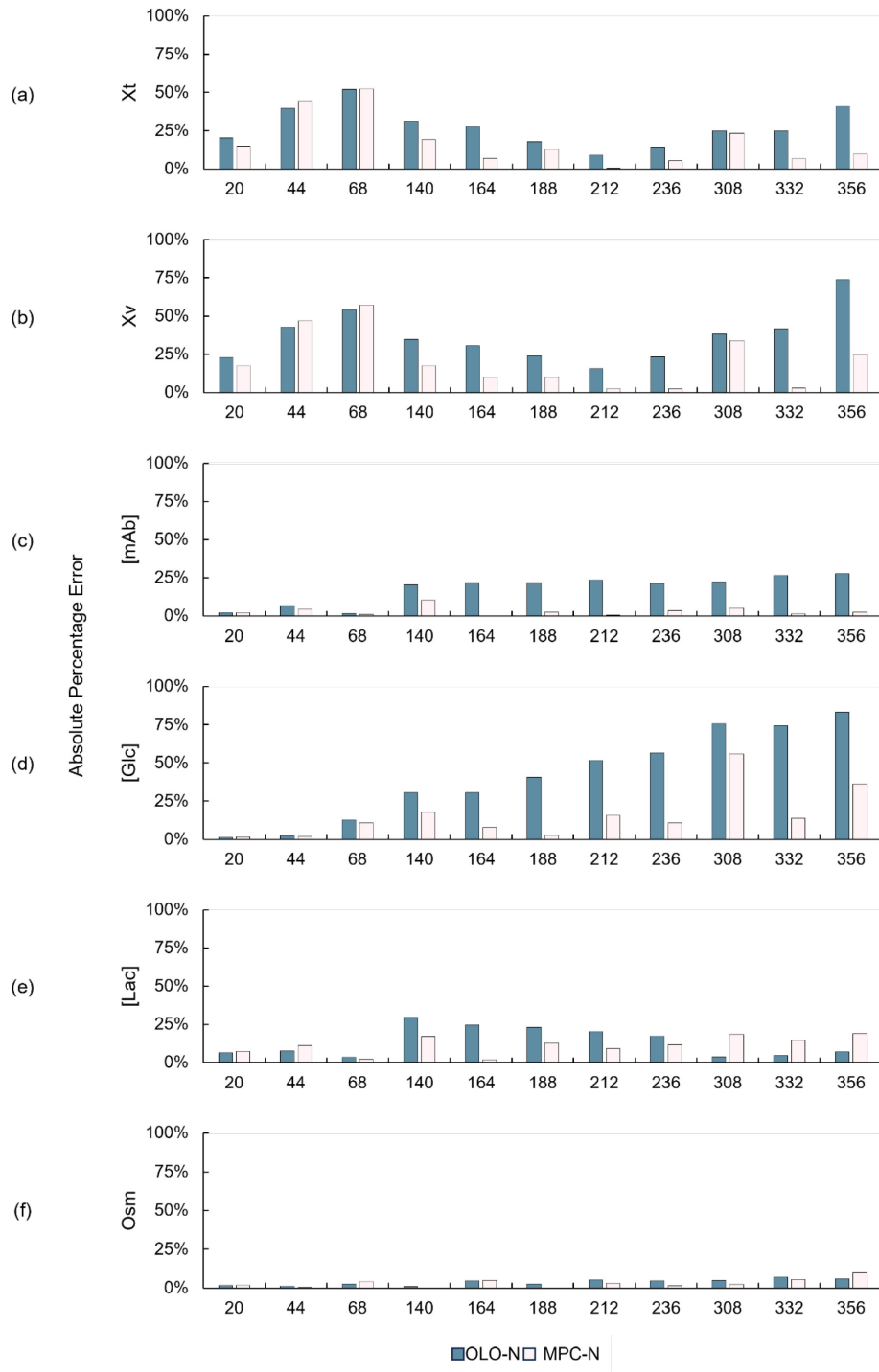


Figure 46 Absolute percentage error between model predictions versus experimental observations when deploying open-loop optimisation (OLO-N) and non-linear model predictive control (MPC-N) during normal operating conditions. Normalised using min-max normalisation.

In Figure 45, the model can track the experimental results of MPC-N more accurately than OLO-N showing the superior benefits of updating the state prior to each control action during MPC. This is especially visible in Figure 46c where the absolute percentage error when tracking [mAb] is close to zero during MPC-N. To further analyse model performance during OLO-N and MPC-N, the R^2 , root mean square error (RMSE) and Pearson correlation coefficients (PCC) were calculated and presented in Table 19. Here the R^2 when tracking glucose is particularly low for both OLO-N and MPC-N, although worse for OLO-N indicating a large plant model mismatch. That said in both cases the R^2 values when tracking [mAb] are close to 1 indicating low plant model mismatch for this variable. A larger RMSE when tracking [mAb] for OLO-N is due to poorer predictions especially at the end of the process as visible in Figure 45.

Table 19 Model performance when deploying open-loop optimisation (OLO-N) and non-linear model predictive control (MPC-N) during normal operating conditions. RMSE: root mean square error in units of measurements for X_t (cells.mL⁻¹), X_v (cells.mL⁻¹), [mAb] (mg.L⁻¹), [Glc] (mM), [Lac] (mM), Osm (mOsmol.Kg⁻¹). PCC: Pearson correlation coefficient.

	OLO-N			MPC-N		
	R2	RMSE	PCC	R2	RMSE	PCC
X_t	0.97	2.56E+06	0.99	0.94	1.50E+06	0.97
X_v	0.94	3.46E+06	0.97	0.91	1.85E+06	0.96
[mAb]	1.00	297.03	1.00	1.00	45.69	1.00
[Glc]	0.23	29.54	-0.48	0.55	5.52	0.74
[Lac]	0.95	5.52	0.98	0.98	4.09	0.99
Osm	0.98	17.44	0.99	0.80	15.00	0.90

5.4.4 Experimental Observations vs In-Silico Optimisation

Optimisation was also performed prior to operation for all processes. As the objective function was an economic objective, continuously minimising the endpoints only, the endpoint values of the controlled variables obtained during in-silico optimisation (Section 5.3.3) can be used as a target setpoint for experimental demonstrations. Specifically, the percentage error (PE) between the anticipated optimal endpoint values of each controlled variable in the objective function (Eq. 29 a) and the experimental observations is calculated and shown in Figure 47. For reference, the initial model-building experiments (MF-N, MF-L, and MF-H) were also compared to the in-silico optimal results and are displayed solely as a relative indication of

what is achievable experimentally. In addition, the mean percentage error (MPE, %) is also shown in Figure 47. Finally using Eq. 30 a weighted mean percentage error (WMPE, %) is calculated. Here, a higher product titre and IVCD are rewarded, whereas a higher IntLac, IntOsm, and total glucose feed is penalised similar to the objective function presented in Eq. 29 a.

$$WMPE = [mAb]_{PE} + IVCD_{PE} - IntLac_{PE} - IntOsm_{PE} - Total\ Glucose\ Feed_{PE} \quad Eq. 30$$

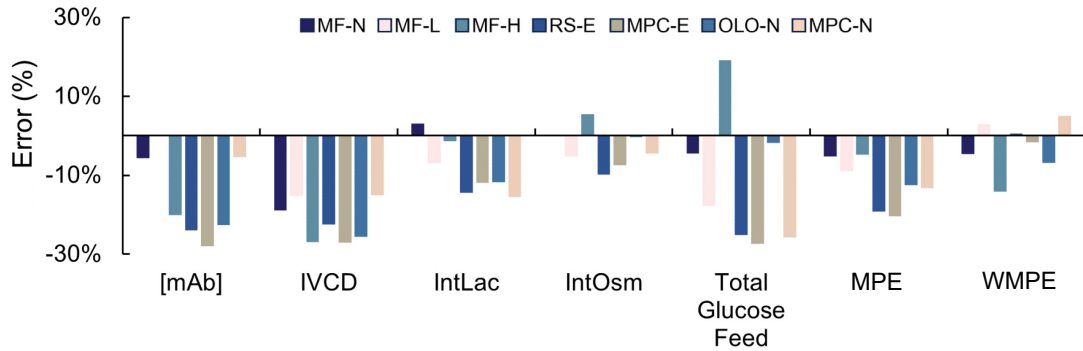


Figure 47 Percentage error between the anticipated optimal performance and achieved performance of variables within the objective function. MF-N: Manual feeding a nominal amount of glucose. MF-L: Manual feeding a low amount of glucose. MF-H: Manual feeding a low amount of glucose. RS-E: reactive scheduling in the presence of a major operating error. MPC-E: model predictive control in the presence of a major operating error. OLO-N: open-loop optimisation experiment in the absence of major operating errors. MPC-N: model predictive control in the absence of major operating errors.

5.4.5 Comparison of Variables' Endpoint Value

A comparison is shown in Figure 48 between the endpoint values of experimentally observed variables during all process control strategies including manual glucose feeding (MF-N, MF-L, MF-H), reactive scheduling during an operational error (RS-E), model predictive control during an operational error (MPC-E), open-loop optimisation during normal operation (OLO-N), and model predictive control during normal operation (MPC-N). Observations are described in-depth during the following discussion.

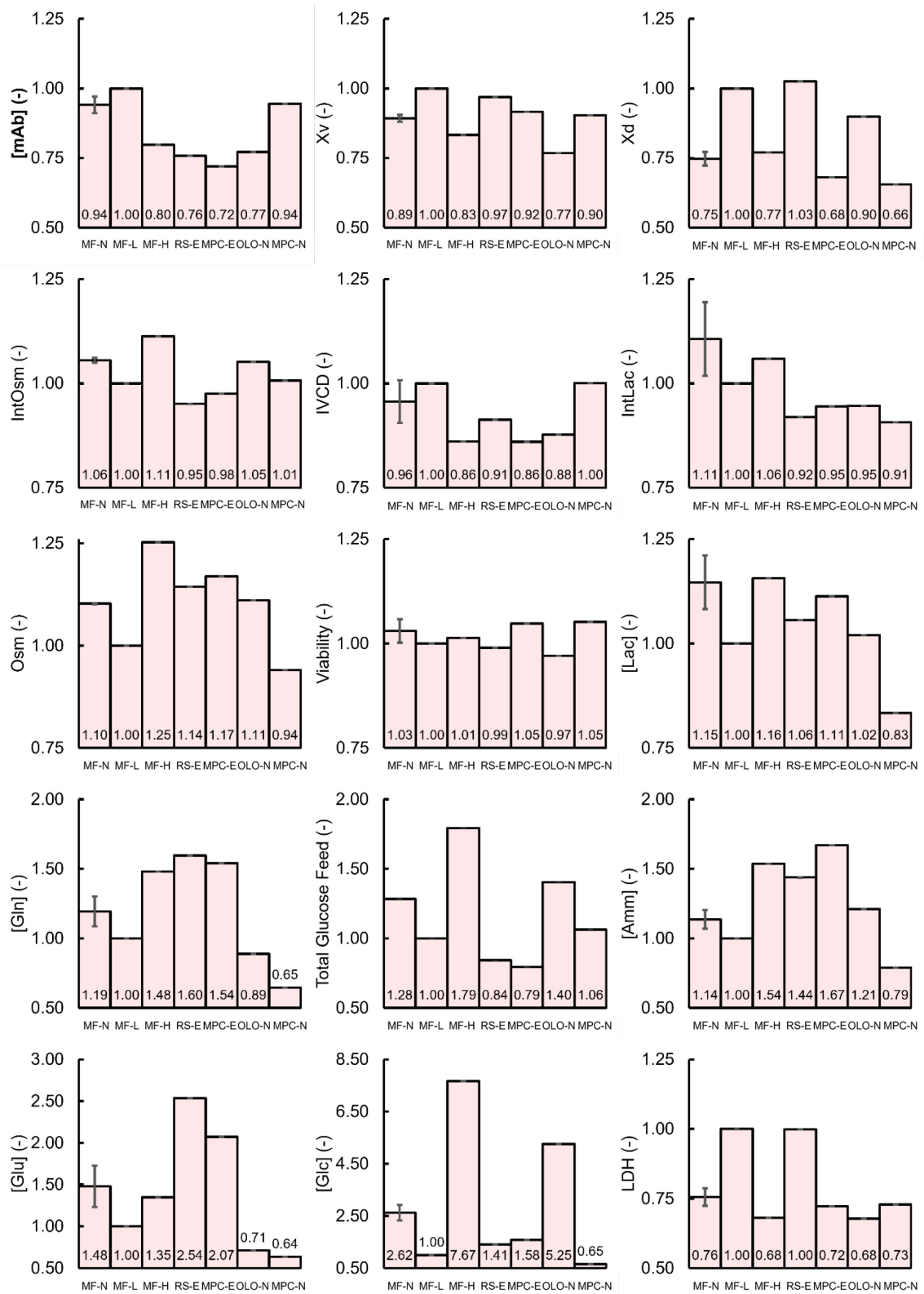


Figure 48 Comparison of all endpoint values of variables measured during cell culture experiments normalised against the manual feeding experiment with the highest product titre, MF-L.

5.5 Discussion

Several computational tools were used to derive a suitable operating strategy including *in silico* input-output analysis and sensitivity analysis. From this, an objective function was defined and paired with suitable constraints. *In silico* optimisation was performed producing predictions of the optimum trajectories of the modelled states. Following this, 2L scale bioreactors were used for cell culturing and to demonstrate several novel process control strategies manipulating the volume of glucose fed.

Experiment RS-E utilised reactive scheduling, a process control algorithm that was demonstrated previously on CHO cell cultures using *in silico* demonstrations to control glycosylation (Kotidis et al., 2021). However, this research is the first to test reactive scheduling experimentally on a fed-batch CHO cell culture, albeit attempting to prevent batch loss due to an operational error. Here, reactive scheduling was used to optimise a multi-variable objective function directly after an operational error to maximise product titre, and IVCD whilst minimising integral extracellular lactate concentration, integral osmolality, and total feed volume of glucose. Reactive scheduling (RS-E) was also compared to a non-linear model predictive control strategy during an operational error (MPC-E). In this instance, the same objective function was used as in RS-E albeit optimisation was performed prior to every control action. A drawback here is that RS-E and MPC-E were not compared to manual glucose feed strategy during the presence of an operational error, thus there is scope for this in future research. Finally, open-loop optimisation (OLO-N) was compared to non-linear model predictive control (MPC-N) during normal operation.

Model predictive control has been used to control fed-batch CHO cell cultures experimentally previously (Aehle et al., 2012; Craven et al., 2014; Jones et al., 2023; Rashedi et al., 2023a; Sarna et al., 2023; Schmitt et al., 2019; Zupke et al., 2015). In addition, mechanistic models have previously been deployed within a model predictive control strategy for fed-batch CHO cell cultures and demonstrated *in silico* (Ohkubo et al., 2023) and experimentally (Aehle et al., 2012; Craven et al., 2014; Zupke et al., 2015). Of the latter experimental demonstrations, these were used to control oxygen consumption (Aehle et al., 2012), extracellular glucose concentration (Craven et al., 2014), and high-mannose N-linked glycosylation (Zupke et al.,

2015). However, this research represents one of the first studies to experimentally demonstrate a mechanistic model within a model predictive control strategy to directly control a key performance indicator, product titre during experiment MPC-N. Here, several other variables are controlled jointly within an economic objective function whereby only the endpoint is minimised within a receding control and prediction horizon.

In general, caution should be taken when comparing experiments directly as it remains challenging to precisely replicate conditions batch-to-batch. Further, the complexity of the process can make operation challenging as evident during RS-E and MPC-E where the operational error occurred resulting in the unintentional addition of media. This must be considered when comparing RS-E and MPC-E to other experiments.

5.5.1 Comparing Variations in Manipulated Variables

Overall, when comparing the deviations in manipulated variables across experiments (shown in Figure 41 and Figure 42), the cell culture demonstrates robustness to variations in glucose feeding strategy as shown in Figure 43 and Figure 45. Experiments RS-E and MPC-E feed the lowest amount of glucose (Figure 42) primarily due to the lower requirements of a predominantly lower viable cell density occurring throughout the process versus other cell cultures. This was caused by the operational error. This operational error meant that compared to the highest IVCD achieved (in MF-L and MPC-N) across all experiments, a 9% lower IVCD was achieved using reactive scheduling in the presence of an operational error (RS-E), and a 14% lower IVCD was achieved using model predictive control in the presence of an operational error (MPC-E). Out of the feeding strategies that utilised optimisation (RS-E, MPC-E, OLO-N, and MPC-N), open-loop optimisation (OLO-N) fed the highest volume of glucose (Figure 42). This is due to an over-prediction of cell density during open-loop optimisation (OLO-N). For example, it was previously discussed (in Chapter 4) how the model offers poor prediction capability during late-stage cell culture. Recent research has focused on improved prediction of late-stage cell culture using model-building tools (Okamura et al., 2022). Here, however, the model is utilised online and when using model predictive control, one-step ahead predictions are used to improve prediction capability. For example, MPC-N demonstrated how model predictive control can be used to alleviate such plant model mismatch online. This is showcased

in MPC-N compared to OLO-N where there was a 46% reduction in RMSE on the prediction of viable cell density (Table 19).

5.5.2 Comparing Reactive Scheduling and NLMPC during an Operational Error

Initially, reactive scheduling (RS-E) was operated alongside non-linear model predictive (MPC-E) control during a severe operating error. Following an operating error, the experiments were continued to ascertain the degree to which the process can be recovered when using model predictive control. Looking at Figure 32, there is only a minor difference between the viable cell density achieved during the experiments testing each process control strategy. Here, RS-E achieved a 6% higher peak viable cell density than MPC-E. Further, the endpoint viable cell density of RS-E is 5% higher than MPC-E shown in Figure 37. There is a slightly greater difference in total cell density. In detail, the peak total cell density of RS-E is 11% higher than MPC-E whilst the endpoint total cell density of RS-E is 11% higher than MPC-E (Figure 32). These variations could point towards slightly reduced growth during late-stage cell culture or slightly increased cell death as indicated by a 51% higher endpoint dead cell density in RS-E (Figure 37). Further, in RS-E, there is a 38% greater level of LDH than in MPC-E. This could be an indicator of a reduced amount of cell lysis in RS-E (Klein et al., 2015). However, much of this may solely be due to changes in cell density rather than differences in cell death rate as there is only a 6% lower endpoint viability in RS-E compared to MPC-E shown in Figure 37. The differences in viability could be explained by the differing process control strategies. In detail, looking at Figure 38, there is only a minor difference between the total feed volume of glucose in RS-E against MPC-E. However, a large difference occurs on day 13 (Figure 30), where model predictive control (MPC-E) reacts to the over-prediction of cell density and thus feeds a volume of glucose close to zero relative to reactive scheduling (RS-E) which only reoptimized after an operational error occurred on day 2. On the contrary, the day 13 feeding during RS-E causes a large increase in osmolality on day 14 as shown in Figure 32. Here despite RS-E having a 20% higher measured viable cell density, both RS-E and MPC-E were at a similar osmolality within 5 mOsm.kgH₂O⁻¹ of hyper-osmotic conditions. This coincides with the slight decrease in viable cell density after day 13, which may have contributed to decreased

viability, especially in RS-E where higher cell densities occurred at close to equal, hyper-osmotic conditions also present in MPC-E.

The extracellular concentration of glucose in RS-E is lower than that of MPC-E (Figure 43) throughout the cell culture. This is especially the case on day 13 where there is a large plant model mismatch, and the observed extracellular concentration of glucose is greater than the predicted extracellular concentration of glucose. This may be caused by the over-prediction of viable cell density therefore over-feeding of glucose on day 10. Such an over-prediction may be due to the poor prediction of osmolality (Figure 43). Here osmolality is derived via an algebraic equation in the process model in a similar way to prior studies (Alhuthali & Kontoravdi, 2022), crucially not as a differential equation. As such, its value is estimated based on the value of other process variables. Therefore, despite osmolality being routinely measured during cell cultures e.g., using freezing point depression (F. Li et al., 2010), in this model the value for osmolality cannot be directly reinitialised prior to optimisation as is visible in Figure 43. This may be suitable under the normal operating conditions which the model was structured and parameterised within; however, it is clearly a weak estimate of osmolality in this experiment. Here, the operational error causes excess media to be added to the cell culture. The media consisted of many different components of unmeasured osmolality, each without their own mass balances within the model nor included within the prediction of osmolality. Therefore, changes in osmolality due to media addition were not accounted for in the model. Because the growth rate is calculated using osmolality as an inhibitory variable, inaccurate predictions of osmolality also affect the viable and total cell density as is evident particularly in late-stage cell culture.

Across MPC-E and RS-E, other process variables are predicted relatively accurately (Figure 43). For example, although there is a large absolute percentage error close to 50% in the prediction of extracellular lactate concentration at 47 hours shown in Figure 44, the remainder of the process sees lower errors and an R^2 close to 1. Product titre is also predicted relatively accurately. However, reactive scheduling (RS-E) has a 5% higher product titre than model predictive control achieves (MPC-E) despite different process control algorithms being used. This is primarily a consequence of RS-E achieving a higher viable cell density and higher IVCD than MPC-E. In addition, and as mentioned, in the presence of an operational error the

model is extrapolating into conditions which it was not structured or parameterised within. As such, an under-performance is expected. In other recent research where control of product titre is attempted, there is also a similar difference (1.10-15.20%) in endpoint product titre between process control algorithms (Rashedi et al., 2023a). Although the effort required to implement model-predictive control at a manufacturing level may be greater than reactive scheduling, reactive scheduling inherently requires a defined definition or awareness of when a process deviation occurs. Therefore, there is a requirement in future research for a cost-benefit analysis.

5.5.3 Comparing Open-loop Optimisation and NLMPC during Normal Operation

Under normal operation, open-loop optimisation was compared to model predictive control with predicted trajectories shown in Figure 45. Here there is a visible difference between the total cell densities achieved in each control strategy. Notably, model predictive control (MPC-N) achieves a 16% and 8% higher peak total cell density and endpoint total cell density, respectively, than open-loop optimisation (OLO-N). Similarly, model predictive control achieves a 20% and 15% higher peak viable cell density and endpoint viable cell density, respectively, than open-loop optimisation (OLO-N). Further, Figure 45 demonstrates how model predictive control continuously updates the value of viable and total cell density and as such continuously corrects the feeding strategy to account for the most recent measurements. The beneficial effects of this model correction are visible when comparing the extracellular glucose concentration of OLO-N and MPC-N. Here, it is apparent that model predictive control (MPC-N) continuously manipulates glucose feeding to maintain glucose concentrations closer to the predicted optimum extracellular glucose concentration trajectory. On the contrary, for OLO-N, the extracellular concentration of glucose continuously increases leading to continuous over-feeding and increased plant model mismatch visible in Figure 46d. The over-feeding of glucose increases osmolality (Figure 45), which has been shown to inhibit cell growth (see Chapter 4 and previous research (Alhuthali et al., 2021)).

Glucose over-feeding is also known to promote the production of inhibitory metabolites (Schmitt et al., 2019). Here, the peak extracellular lactate concentration during OLO-N was 13% higher than MPC-N (Figure 45) demonstrating increased lactate production possibly due

to over-feeding of glucose during open-loop optimisation (OLO-N). Further, the endpoint point extracellular lactate concentration during model predictive control was 18% lower than during open-loop optimisation (Figure 45). Meanwhile, the decrease in the extracellular concentration of lactate from the peak concentration to the endpoint during model predictive control (MPC-N) concentration was 29% greater than during open-loop optimisation indicating increased lactate consumption (Figure 45). The unfavourable lactate trajectories observed when using open-loop optimisation could be due to glucose being converted to pyruvate during glycolysis, where pyruvate is not utilised and subsequently converted to lactate. However, a full flux balance analysis is required to verify this. In general, these results are in agreement with previous research and show the necessity to tightly control glucose thus improving lactate production and consumption. Albeit lactate production and consumption are also influenced by the cell line in addition to process conditions (Gagnon et al., 2011; Matthews et al., 2016). Despite this, there are few examples in open research that explicitly control lactate using model predictive control (i.e., including lactate within the objective function) during fed-batch CHO cell culturing. Recent examples manipulated pH to solely control lactate (Schmitt et al., 2019). However, in the present research, lactate was directly controlled by including the integral extracellular lactate concentration within the objective function in addition to several other variables. The robustness in achieving a high product titre and improvements in reductions in extracellular lactate concentration may be due to this novel approach. However, testing on different cell lines and conditions is necessary to further validate this strategy.

Finally, comparing the product titre in Figure 45 there is a notable difference between the performance observed during open-loop optimisation (OLO-N) and closed-loop model predictive control (MPC-N). The implementation of model predictive control results in a 22% increase in product titre versus open-loop optimisation. Here, it can be deciphered that a correctly controlled glucose feeding strategy, results in a higher viable cell density thus a more productive cell culture. For OLO-N, the absolute percentage errors observed when for product titre are far greater than for MPC-N indicating a large plant model mismatch (Figure 46c). Further, in Table 19 the RMSEs for MPC-N are always lower than for OLO-N indicating a lower plant model mismatch. Overall, although this improvement in a key performance indicator for a highly complex non-linear process represents a significant achievement, additional research would be favourable to scope the limitations of model predictive control.

This is especially the case as model predictive control did not perform favourably during an operational error compared to other process control strategies. That said, this represents an exciting opportunity to explore closing the loop on previous studies that utilised open-loop optimisation (Kotidis, Jedrzejewski, et al., 2019) and ascertain what improvements if any, model predictive control offers.

5.5.4 Comparison of In-Silico Optimisation vs Experimental Observations

Although the performance observed when using each process control strategy is important, the ability to correctly control the system to a desired optimal setpoint is equally important. A key challenge is doing so in the presence of system uncertainty. Such uncertainties are notable in fed-batch CHO cell cultures, for example, those derived from parametric uncertainty, minor perturbations of initial conditions (e.g., due to differences during the seed train), or disturbances (e.g., anti-foam or base additions). In this research, in-silico optimisation is used to derive a target setpoint which is compared to experimental observations using percentage errors. It is clear from Figure 47 that OLO-N and MPC-N have a mean percentage error closer to zero versus RS-E and MPC-E. This is expected as OLO-N and MPC-N are operating under normal conditions whereas, in RS-E and MPC-E, an operational error occurred significantly affecting process operation. When considering the weighted mean percentage error, MPC-N performs best, followed by RS-E both of which remain the only positive weighted mean percentage errors (aside from MF-L). Open-loop optimisation (OLO-N) performs poorly using this metric although this is primarily because of the relatively high volume of glucose fed compared to RS-E, MPC-E, and MPC-N. When comparing reactive scheduling (RS-E) versus model predictive control (MPC-E) during the presence of an operational error, the benefit of model predictive control here is unclear where similar weighted mean percentage errors are achieved, 0.56% and MPC-E -1.63% respectively. Looking specifically at $[mAb]_{PE}$, for MPC-N this is -5.33% whereas OLO-N has a $[mAb]_{PE}$ of -22.64%. This represents a 76.46% decrease in tracking an optimal setpoint of product titre when using model predictive control (MPC-N) versus open-loop optimisation (OLO-N) during normal operating conditions. Similar trends were observed comparing the $IVCD_{PE}$ during OLO-N and MPC-N of -25.60% and -15.10% respectively. However, during the presence of an operational error MPC-E performs worse than RS-E achieving a lower $[mAb]_{PE}$ and $IVCD_{PE}$, whilst achieving a higher $IntLac_{PE}$ and

IntOsm_{PE}. This reinforces the question of the cost versus benefit when deploying model predictive control over reactive scheduling. Without further experiments, it is challenging to distinguish whether this is due to the model over-extrapolating during an operational error or due to the specific process control technique. In contrast, during normal operation, model predictive control demonstrates an excellent ability to achieve an objective function. Although MF-L achieves a [mAb]_{PE} closer to zero than MPC-N, this is not suitable for direct comparison as MF-N, MF-L, and MF-H were operated without any predefined optimal setpoints. Overall, utilising model predictive control in the current conditions showed clear advantages versus other process control strategies offering a balanced tracking ability across multiple variables. Future experiments should test different weighting strategies and the sensitivity of each weight attached to each variable within the objective function.

5.5.5 Comparison of all Process Control Strategies

In Figure 48, the difference in endpoint product titre between MF-L or MPC-N may be insignificant versus the error observed during MF-L. Despite no improvement observed when using model predictive control versus manual glucose feeding, there is an important distinction in that model predictive control may be used to improve process robustness and consistently achieve desired setpoints across all controlled variables. When comparing MF-L to MPC-N, no improvement in product titre could be because glucose is not significantly influential on the process and instead other metabolites could be manipulated. In addition, this may also be due to the use of a highly developed industrial cell line, designed to be robust. Previous studies also allude that process control strategies can be cell line dependent where a closed-loop feeding strategy showed no benefit to the closed-loop control versus a manual feeding strategy (Lu et al., 2013). Meanwhile, other studies primarily focused on batch reproducibility (Aehle et al., 2012) rather than solely maximisation of performance. Aside from this, there is a minor difference in product titre when comparing open-loop optimisation (OLO-N) to model predictive control (MPC-N). Further, there is little difference between the endpoint product titre of RS-E and MPC-E. For RS-E and MPC-E, the lack of apparent difference in product titre highlights the need for further analysis of reactive scheduling under normal operating conditions. This will allow a stronger conclusion to be drawn as to whether the lack of

difference is due to the process control strategies employed or due to the major operational errors.

The increased endpoint product titre in MF-L and MPC-N relative to other process control strategies can be imparted to the large integral viable cell density (Figure 48). That said, when looking at the endpoint viable cell density for MPC-N this is lower than MF-L potentially due to increased cell death or reduced growth during late-stage cell culture as previously discussed. This highlights intuitively that for maximising productivity, integral viable cell density is suitable to include within an objective function for model predictive control as was also used in recent studies (Rashedi et al., 2023b).

Interestingly, the dead cell density is the lowest (Figure 48) whilst viability was the highest out of all process control strategies when deploying model predictive control during an operational error (MPC-E) and model predictive control under normal operation conditions (MPC-N). This is interesting as neither variable was included in the set of controlled variables within the objective function. The low accuracy during measurement of cell densities and low number of repeat experiments means that there isn't a statistically significant difference between viability in MF-L and MPC-E or MF-L and MPC-N. There is a slightly greater difference in the concentration of dead cells. But within the model, the death rate is modelled as a constant variable and as such the model lacks insight into which variables influence this preventing controllability. One possibility is that during MF-L, there were instances of growth limitation for example due to a low extracellular glucose concentration which caused cell death. This is corroborated by MF-L having a higher concentration of lactate dehydrogenase an indicator of cell lysis, and a higher concentration of dead cells compared to MPC-E and MPC-N. With further analysis, it was also found that the minimum experimentally observed extracellular glucose concentration in MF-L was 18.7% lower and 23.53% lower than the minimum experimentally observed extracellular glucose concentration in MPC-E and MPC-N, respectively. However, there may be other growth-limiting nutrients that were not monitored experimentally or included within the model which limits the scope of this analysis.

Looking at the endpoint extracellular lactate and ammonia concentration in Figure 48, there is an evident increase when deploying model predictive control during MPC-E compared to MF-L. As mentioned, it is speculated that this is because the model is extrapolating rather than due

to the process control algorithm. However, there is a slight decrease in the endpoint extracellular lactate concentration during model predictive control (MPC-N) versus manual glucose feeding (MF-L) relative to open-loop optimisation. During other studies that aim to control lactate using model predictive control, trajectories using this were very similar to a manual feeding strategy. In addition, a slightly higher endpoint extracellular lactate concentration was achieved in all process control strategies than when using manual feeding (Rashedi et al., 2023a). However, this could be because, in the aforementioned study, lactate was not explicitly minimised rather the error between a reference trajectory was minimised. Conversely, in the present study, lactate was explicitly minimised whilst a reduction in extracellular ammonia and lactate concentration is noticeable.

5.6 Summary

This chapter develops, presents, and compares several novel process control strategies for fed-batch mammalian cell culture systems. It is one of few works to experimentally demonstrate advanced process control of a key performance indicator, product titre, on CHO cell cultures. Further, there are very few if any studies to do so using a mechanistic model. The advantage of demonstrating this is that mechanistic models have been shown to outperform data-driven modelling strategies during instances of low data availability, a common occurrence especially during early-stage process development. Finally, during experimental validation of the control strategies, only process variables that are readily available using widely deployed process monitoring equipment are required, a key advantage of this work and step towards easing access to advanced process control, and thus automation of upstream biopharmaceutical processes.

Firstly, the development of a process control strategy using model-based tools was also demonstrated showcasing how input-output analysis, sensitivity analysis, and *in silico* optimisation can be utilised. Here a previously built model was engaged that predicts, based on changes to a glucose feeding schedule, several key process variables including product titre, viable cell density, total cell density, extracellular concentration of glucose, extracellular concentration of lactate, and osmolality. From this, an optimisation was proposed including terms within the objective function and constraints. The multi-variable objective function

aimed to maximise product titre and integral viable cell density, minimise inhibitory components including the integral osmolality, the integral extracellular concentration of lactate, and minimise the total volume of glucose fed. Constraints were chosen loosely to ensure feasibility and help the optimisation converge to the true global optimum. This risk is especially prominent when utilising a non-linear model during model predictive control. As such the manipulated variable, the volume of glucose fed at each feed interval, was loosely constrained. In addition, the only other constraints were on the lower bound of the extracellular concentration of glucose to minimise model extrapolation, and the upper bound of osmolality to prevent hyperosmotic conditions.

During experimental studies, the opportunity to demonstrate advanced process control strategies during a severe operating error arose whereby reactive scheduling and model predictive control were deployed. Following this, open-loop optimisation and model predictive control were demonstrated during normal operation. All process control strategies, including manual glucose feeding, were compared.

During instances where the model was deployed online, there are clear examples showcasing how the improved utilisation of available data can allow for improved model prediction capability. In detail, by using one-step ahead predictions plant model mismatch is reduced. Because of this, more informed decision-making can take place. Although in some process variables, in particular product titre, despite an improvement in model predictive control compared to open-loop optimisation, there was no improvement in product titre when comparing model predictive control to manual glucose feeding. However, results show that model predictive control could improve process robustness by accounting for plant model mismatch and achieving close to identical product titre. In addition, when deploying model predictive control, viability was higher, the dead cell concentration was lower, and levels of lactate dehydrogenase were lower, thus cell lysis is assumed to be lower (Chan et al., 2013; Klein et al., 2015; O'Mara et al., 2019), therefore host cell protein content may also be lower (Tait et al., 2012). The effects of this should have marked improvements downstream where reduced purification efforts to achieve the same product quality may be possible. Overall, model predictive control employed an economic multi-variable objective function, and showcased a balanced performance, and the best setpoint tracking performance in terms of the

weighted mean percentage error, therefore paving the way for more ambitious control objectives to be included in the future.

Regarding future research, firstly, these results clearly contrasted the poor performance of the model whilst extrapolating versus when utilised closer to the conditions at which it was structured and modelled within. At present, the model is used as a “digital shadow” solely to predict future outcomes based on changes in the states. Because of the limited data used during model-building, it is challenging to obtain an accurate and robust model. Along with increased experiments, an improvement to the current method may be to use an estimator to predict the process outputs more accurately. Failing this, the model may be used within a “digital twin”. Here the model itself would be passed through an online re-identification algorithm either to update process parameters or model structure based on newly available data. Secondly, to truly evaluate model predictive control strategies, the type of model used should be varied (mechanistic, hybrid, data-driven) whilst keeping all other decisions identical (e.g., objective function, scale etc.). To further improve the robustness of the process control strategy there are several possible paths of continuation. To minimise uncertainty created by perturbations in initial conditions, future research could include the similar deployment of advanced process control within the seed train which in prior process control studies was shown to have a notable impact on final process performance (Gagnon et al., 2011). To minimise uncertainty due to parameter selection, multi-stage optimisation could be utilised within model predictive control whereby rather than a set of parameters defined prior to optimisation, a range of parameters are defined in which the true value may be as well as minimising uncertainties arising. Finally, to minimise uncertainty due to process noise, data smoothing algorithms could be used to alleviate this e.g., Savitzky-Golay filters.

Chapter 6 Conclusions and Further Research

Current process development workflows for fed-batch CHO cell cultures rely heavily on laborious and costly experiments to derive suitable processes and operating conditions. Once these conditions are defined, the predominant method for process control is to use regulatory control of base conditions (e.g., pH, temperature) using predefined setpoints. Importantly, the control of key performance indicators is performed in an open-loop manner e.g., using a predefined batch recipe to achieve a desired product titre. Advances in process modelling of fed-batch CHO cell cultures have opened the opportunity to apply advanced model predictive control strategies. Meanwhile, models can also be used to significantly cut costly and timely experiments by using *in-silico* tools to design improved processes. However, model development remains challenging due to the expense of acquiring data, and in general the highly manual nature of elucidating the process dynamics of fed-batch CHO cell cultures. Another challenge in the implementation of model predictive control relates to the difficulty in access to real-time data. In this respect, there is a general lack of process analytical technology especially that which is available online. During operation, process measurement remains highly intricate, and the following issues are common: (a) missing data, (b) data measurements at differing times, (c) erroneous data due to mis-calibrated instruments, or (d) measurements with large variances. Therefore, when designing model predictive control strategies care must be taken to ensure that any modelled variables are paired with a reliable method of real-time feedback. In general, there are not only few experimental demonstrations of model predictive control applied to fed-batch CHO cell cultures, but even fewer that place focus on the design and justification of the control strategy. In summary, this thesis aims to bridge this gap in the modelling and control of fed-batch CHO cell cultures.

In Chapter 2, a literature review was presented highlighting how automation-enabling technologies could advance process development and manufacturing of biopharmaceuticals.

Specific emphasis was placed on reviewing modelling workflows that exist in literature including mechanistic, hybrid, and data-driven methodologies. The highlights and pitfalls of each were shown with relevant literature. In addition, the importance of tight control of upstream biopharmaceutical processes along with the research gaps present in developing a controller that guarantees process performance and robustness. A key focus was placed on the research in open-loop optimisation versus the limited research in closed-loop control strategies. Further, the effects of process variables on mammalian cell growth and metabolism were discussed with methods from the literature reviewed. Finally, various computational tools were summarised including applications of sensitivity analysis and in-silico optimisation to control key performance indicators including cell growth and productivity. Following Chapter 2, an overview of the materials and methods used in this thesis was given in Chapter 3.

In Chapter 4, a mechanistic modelling workflow was demonstrated using an industrial CHO cell line describing CHO cell growth and metabolism. Initially, three fed-batch CHO cell culture experiments were performed in 2L bioreactors. Following this, an automated script was developed to enable efficient data analysis of experimental data. Future use of this is expected to allow for in-depth real-time data analysis e.g., real-time PCA, or teal-time DoE. A design feature of the workflow presented was in the selection of metabolites and variables to include within the model whilst preventing reliance on variables that are not significantly influential on the process output. In doing so, one can reduce reliance on excessive process monitoring equipment and the minimise logistical efforts attached to building data flows. If process variables are included unnecessarily, this could increase the complexity of the parametrisation problem where more degrees of freedom are introduced that do not significantly influence the output. Following the initial selection of a model structure and set of parameters, sensitivity analysis was deployed to test which parameters could be fixed, and which parameters the model was sensitive to thus may require re-tuning upon receipt of new data. Overall, a key achievement was made in demonstrating the power of mechanistic models during instances of low data availability, using just three experiments to build a model that was later deployed in real-time under experimental conditions different from those which the model was parameterised upon.

In Chapter 5, sensitivity analysis was employed to aid the design of a novel process controller. Here in-silico studies were performed each with varying glucose concentrations added on each of the days of the culture. The bounds for the sensitivity analysis were set such that the maximum volume of glucose fed was within the conditions which the model was validated within, whilst also preventing over-feeding of glucose by setting the minimum volume to zero. The results of this sensitivity analysis allowed for insight into which variables may be challenging to control in relation to each respective day of glucose feeding. It was revealed that product titre and cell concentration appeared to be particularly challenging to control during late-stage cell culture. In addition to the sensitivity analysis, in-silico dynamic optimisation was also performed by testing a range of objective functions. Ultimately, a multi-objective function was chosen on the basis that this will guide the process to success across multiple variables, whilst also providing the opportunity to simultaneously achieve high product titre, reduce by-product accumulation, and minimise inputs. The final objective function was then implemented into a model-predictive controller and tested experimentally at the same scale as the model-building experiments. Here, the process inputs were determined in real time in a closed-loop manner. Initially, an operating error occurred that caused a significant process deviation. This allowed the opportunity to test whether an advanced process control strategy can recover a process and prevent complete batch failure. However, in this case, it was difficult to ascertain the benefit of model predictive control, albeit this comparison was prone to limitations discussed below. Following this, a successful experiment was performed where here model predictive control demonstrated superior performance against open-loop optimisation, showing robustness with respect to product titre, increased viability, and a reduction of by-products compared to manual feeding and open-loop optimisation.

6.1 Main Contributions

Across industry, there is a growing interest towards enabling process automation. However, the challenges in modelling and control of fed-batch mammalian cell cultures inhibit progress towards this goal. The overall motivation of this thesis was to advance methods of process modelling and control of mammalian-based cell cultures to enable progress towards process automation within upstream biopharmaceutical processing. To that end, this thesis has

demonstrated how modelling and advanced process control may allow for simultaneous improved process robustness and performance. The contributions of the work presented herein and relevant conclusions are listed below:

1. Three different manual glucose feeding methods were tested experimentally to compare the effects on a fed-batch CHO cell culture with respect to product titre, by-product formation, growth, and productivity.
2. A mechanistic modelling workflow was demonstrated to produce a model describing CHO cell growth, metabolism, and osmolality. Relative to many existing modelling workflows in literature, here a conscious effort was made to ensure the model could be deployed in real-time. To do so, a simplified metabolism was formulated by defining logic to minimise the number of metabolites modelled thus reducing process monitoring requirements. Specifically, all modelled process variables can easily be measured using widely available technologies whilst a key performance indicator, product titre, can be predicted with suitable accuracy. The modelling framework demonstrated may be a valuable tool towards future automated modelling of mammalian cell cultures. Building models more efficiently could enable computational methods such as process optimisation to be deployed more easily for process development whilst reducing reliance on extensive lab-based methods. In doing so there is an opportunity for reduced cost or faster process development timelines.
3. Sensitivity analysis was used to predict how product titre, cell viability, osmolality, and extracellular concentration of key metabolites are affected by the variation in volume of glucose fed according to a fed-batch schedule. In this thesis, results showed that the product titre and cell density were not sensitive to variations in glucose feeding strategy during late-stage cell culture. However, extracellular lactate concentration and osmolality were sensitive to variation in glucose feeding across the cell culture, both of which were previously linked to reduced product titre. As such, the latter two variables were also included as controlled variables. A similar workflow can be used to develop increased process understanding, increased understanding into the behaviour of the model, and in general for the design of model-based control strategies.
4. In-silico model-based dynamic optimisation was performed to test a series of objective functions to maximise product titre, maximise cell density, minimise by-products,

prevent hyper-osmotic conditions, and/or minimise the amount of feed. This can be used in future to predict process performance.

5. A novel method of minimising lactate formation during fed-batch cell culturing was proposed whereby the integral of the extracellular concentration was minimised. This was performed as solely minimising end-point extracellular concentrations of lactate may not prevent exposure to excessive concentration during the process, although this remains to be thoroughly validated. Meanwhile minimising the extracellular concentration of by-products at the next time step may inadvertently reduce growth, although again, this remains to be thoroughly validated. The endpoint value of the integral of the extracellular concentration was instead proposed as a measure of overall exposure to lactate. This same method was also applied to reduce exposure to osmolality throughout the course of the process. Similar approaches could be used to minimise exposure to other by-products instead of solely minimising the measured concentration.
6. For the first time in literature, a mechanistic model was deployed experimentally within a model predictive controller for closed-loop control of product titre during fed-batch CHO cell culturing. In this thesis, it is shown how mechanistic models can be developed during instances of low availability, and as such model predictive controllers can also be developed during low data availability. This is expected to be especially useful when not only predicting the performance but also controlling fed-batch CHO cell cultures during operation in new process conditions.
7. For the first time in literature, model predictive control was used experimentally for closed-loop control to minimise exposure to extracellular lactate concentration and simultaneously control product titre during fed-batch CHO cell culturing. In addition, exposure to excessive osmolality was also controlled whilst process inputs were minimised. This thesis shows how advanced process control could be utilised to simultaneously achieve multiple benefits. For example, minimising by-products may lead to secondary effects e.g., improving product quality or reducing downstream processing costs. In addition, simultaneously minimising inputs could lead to the inclusion of sustainability metrics and thus more sustainable processes with reduced costs economically and environmentally.

8. Open-loop optimisation was tested experimentally using the same objective function as used during model-predictive control. Here a set of control inputs was defined prior to process operation and during normal operating conditions performed worse than using closed-loop control. This could indicate that there are further improvements to be made using closed-loop control specifically where previous research has solely tested open-loop optimisation.
9. Finally, open-loop optimisation, manual feeding, and closed-loop model predictive control were all compared following experimental demonstration during normal operating conditions. Further, an operational error allowed the opportunity to test the limits of advanced process control. This is one of few studies to compare advanced process control experimentally to such a range of control strategies on the same fed-batch CHO cell culture. Further, the underperformance of advanced process control during an operational error relative to other control methods demonstrates a clear need for further research to ensure that, given the potential cost of implementation of model predictive control, there is a clear, consistent benefit.

6.2 Limitations & Future Research

The work within this thesis exhibits limitations thus an opportunity for further research avenues. These can be segmented into (a) improvements to the mechanistic model and modelling workflow, and (b) improvements to the process controller including process control strategy, and process control algorithm. In the following section, these are presented alongside additional ideas for improvements.

6.2.1 Improvements to the Mechanistic Model

An ambitious overall aim would be to move towards automated mechanistic modelling and/or highly generalisable models. However, this requires significant advances in both understanding and the development of mechanistic modelling workflows. In this thesis, a workflow was demonstrated resulting in a model that accurately predicts a fed-batch CHO cell culture within a fixed set of conditions. Only the logic to derive a growth rate equation was showcased whereby limits were set on linear regression parameters to aid the selection of inhibitory and

limiting metabolites. Even so, this alone failed to identify a growth rate equation using the measured metabolites alone. Instead, previous expertise was drawn upon to include osmolality within the growth rate equation. In part, this was due to an obvious limitation linked to the lack of process data, especially of other potentially influential metabolites.

Challenges in obtaining data during mammalian cell culturing are common and include (1) a lack of different experiments performed, (2) a lack of repeat experiments performed, and (3) a lack of process measurements obtained e.g., of other metabolites. On the first point, in other process industries, it is common that when building process models, experimental data is segmented into training data, testing data and validation data. However, in this thesis, the lack of experiments performed meant the data was not segmented and instead the model was assumed to be valid based on its fit to experimental training data. For mammalian cell culturing, this is not uncommon and is in part because of the large costs and efforts associated with experiments. Such costs are particularly prohibitive past flask scale including those within this thesis performed in a single-use bioreactor. Here, the preparation required for running a single set of experiments extends well beyond the batch period which in this thesis was fixed at 15 days. For example, the main goal prior to operation is to prepare a sufficient seeding density which in this thesis was sourced using flasks and rocked bag reactors. Further, complexities associated with the hardware can also increase with scale and require increased preparation e.g., to perform sterilisation. As scale increases, the potential cost and time required for the inoculum or seed train may also increase. With this, there may be more opportunities for contaminations to delay production, albeit with increased scale comes increased safety measures. To prevent complete batch loss, sub-cultures were scaled up in parallel rather than in series and combined once contamination was ruled out. That said, this can increase process costs due to the requirement for more equipment.

Overall, in this work, the lack of experiments was addressed using a mechanistic model which was able to predict the experimental data with sufficient accuracy. Aside from performing increased experiments, future work could segment individual data points within experiments into training data, testing data and validation data. Here, care should be taken to prevent bias. For example, results within this thesis indicate that the prediction of viable cell density during the death phase is more challenging than during the growth phase. With this, validation using

data taken only during the growth phase could mislead any corresponding assessment of model validity across the entire cell culture. To automate mechanistic modelling, a generalisable and fair method of data segmentation should be built into any future automated workflow. Importantly, the kinetic model developed in this thesis was only for a specific cell line, utilised during fed-batch cell culturing, with variation only in the volume of a fixed concentration of glucose further constrained to a fixed feeding schedule. To gain a better understanding of the cell behaviour and produce a more generalisable model, a wider range of process conditions should be tested. Alternatively, rather than producing a single highly generalisable model, adept at predictions across all conditions using a single structure and set of parameters, it may be more suitable to have a set of model structures or parameters that are dependent on process conditions. However, the goal of developing a generalisable modelling workflow would still apply and help reduce model development timelines. In particular, the latter would be widely useful given the variation in cell metabolism that can occur across CHO cell lines.

As well as the lack of different experiments, the limited opportunity to perform experiment repeats inhibits thorough analysis to determine if each experiment was statistically significantly different. For example, the calculation of p-values is inherently sensitive to sample size therefore the more repeats that are conducted the greater the likelihood that significance, or lack of it, can be confidently characterised. Clearly, as was the case in this thesis, this must be mitigated against the risk of insufficient experimental exploration and testing considering the associated high costs. Moving forward, statistical significance should be included within an automated modelling workflow to determine whether experimental training data is statistically significant between experiments. This can be combined with statistical measures presented in this study to identify differences between experimental training data and model predictions including R^2 , PCC, and RMSE. Finally, improvements and wider deployment of PAT could also lead to more accurate measurements of process variables, enabling reduced measurement variance and allowing differences between experiments to be more easily observed, mechanisms elucidated, and subsequently modelled with higher confidence.

With increased process measurement capabilities also comes the opportunity to improve modelling capabilities, model validation, and process control development. For example, in this thesis, only widely deployed and available measurements were taken that included just five

metabolites. As such, it is accepted this is likely to approximate the cell metabolism only loosely. Instead, improved PAT may lead to more frequently available process measurements. The benefits of this could include (1) an increased ability to validate process mechanisms, (2) an increased ability to build more accurate or generalisable process models, and (3) an increased ability to design more effective operating strategies. Aside from improved PAT, steps to validate the mechanisms within the model could include using flux balance analysis. In doing so, the role glucose plays across the cell's metabolism during the specific process operated within this thesis could be revealed and validated against the existing mechanistic model.

Apart from the limitations derived from experimental data, or lack of, there is room to explore the model structure further. In this thesis, a mechanistic model was built prioritising its ease of online application via the careful selection of major and minor metabolites whilst also aiming to prevent over-parameterisation. However, the development of a hybrid structure could also be compared. Hybrid modelling is growing increasingly popular throughout literature and can offer superior performance in predicting lesser-known mechanisms. A major drawback is that data-driven aspects of process models generally require more extensive datasets to train whereas mechanistic models can be built from few experiments as shown in this thesis. As such, this might be more relevant for industrial research. Even so, an automated modelling workflow could implement logic for structure selection that could include mechanistic equations as well as data-driven aspects. In addition, there has been interest in other process industries in probabilistic modelling approaches where model outputs are represented by a distribution e.g., using Gaussian processes. This could be an interesting avenue to explore for CHO cell culturing due to the high degree of uncertainty that exists in the process. Bringing this together, an automated modelling methodology would not be constrained to only mechanistic modelling or only hybrid models, there are instances where each is appropriate or feasible.

The selection of model parameters represents one of the major challenges during mechanistic modelling. Careful systems analysis is required to ensure parameters are biologically relevant. Meanwhile, these can vary depending on the cell line and process conditions. In addition, the parameterisation problem can be highly non-linear and non-convex making the search even more challenging, with a risk of falling on a set of parameters that are only a local optimum.

Finally, the effects of overparameterization are well described in literature. In this thesis, a warm start was used for parameters to drive them towards their true, yet uncertain value. Further, the number of parameters was minimised indirectly via careful selection of only important metabolites, which also meant a reduced amount of process sensor technology. However, further research could focus on integrating the explicit minimisation of parameters within an automated workflow for mechanistic models.

Lastly, with regard to the process model, there has been significant interest across process systems engineering in the development of digital twins. At present, the model is used as a “digital shadow” solely to predict future outcomes based on changes in the states. However, an improvement is anticipated if the model were to be used as a “digital twin” with self-tuning. Here the model itself would be passed through an online re-identification algorithm either to update process parameters or model structure based on newly available data. This could be especially promising in instances where the model has been designed with low data to begin with, but an advanced process control method is required. Here model robustness could increase simultaneously while the process is being controlled. The challenge here will be producing a generalisable workflow that will contain suitable model validity checks for the updated parameter set and/or structure.

6.2.2 Improvements to the Process Controller

Given that as far as the author is aware, this is the first application of a mechanistic model within a model predictive controller to control productivity and other variables during fed-batch CHO cell culturing, there is a significant amount of further exploration possible when using advanced process control for CHO cell culturing. Clearly, further experiments need to be performed to increase confidence in the current results. A similar discussion was made above regarding more data required to further determine the statistical significance, or lack of, of the various process controllers presented in this thesis. Another intuitive limitation is that in this thesis, there was rigidity in the feeding schedule available. This includes the frequency of feed during the feeding schedule which was set to a maximum of once a day to allow for a fair comparison to the current status quo in industry, manual bolus feeding of glucose. In addition, the amount of feed inputs available to manipulate was limited to only glucose. Although this

greatly simplifies the modelling process and implementation experimentally, it also greatly limits the ability to influence the process. This could explain why no significant increases in product titre were made compared to manual control, rather the controller enabled robustness in meeting the maximum achieved product titre. At present, the remaining feed components are generally fed in a pre-made feed of fixed composition. In future work, it is of great interest to manipulate the individual components within for control of the process. A challenge would certainly be to ensure the model could describe the addition of any of the other components that aren't otherwise fixed. Similarly, also ensuring the process analytical technology available was appropriate to provide feedback for advanced process control.

Elsewhere, to truly evaluate model predictive control strategies, the type of model used should be varied (mechanistic, hybrid, data-driven) whilst keeping all other decisions identical (e.g., objective function, scale etc.). This could unlock further improvements in control, especially if more accurate or robust models were tested. But it also presents the question of what the minimum model performance required is for it to be deployed within a model predictive control strategy. This is especially the case if more frequent process measurements and more frequent control actions could be taken and therefore shorter prediction horizons were required. Alternatively, models that utilise all data more effectively should be considered. For example, recurrent neural networks could be used to predict process outputs using the full historical time series available rather than solely the most recent measurement.

In general, fed-batch CHO cell cultures exhibit large degrees of uncertainty within the process. There are a range of methods that could be tested in the future to enhance the proposed methods in this thesis. To minimise uncertainty created by perturbations in initial conditions, future research could include the similar deployment of advanced process control within the seed train which in prior process control studies was shown to have a notable impact on final process performance (Gagnon et al., 2011). To minimise uncertainty due to parameter selection, multi-stage optimisation could be utilised within model predictive control whereby rather than a set of single parameters being defined prior to optimisation, a set of parameter ranges are defined within which the true value may be. Elsewhere, to minimise uncertainty for example stemming from instrumentation error, data smoothing and soft sensor algorithms could be used to

alleviate this such as Savitzky-Golay and Kalman filters. This may allow for improved controller performance if the controller is informed by a truer value.

Other research could focus on the objective function, in particular testing whether the in-silico dynamic optimisation results for each scenario tested within this thesis are reflected experimentally. Further variations could also include having a time-dependent objective. For example, during the lag and log phase of cell growth, growth could be maximised. Following this in the stationary phase of the cell culture, productivity could be maximised whilst cell death is minimised to extend the batch duration. Alternatively, batch duration could be directly optimised to maintain a productive cell culture for longer or reach a target product faster. Finally, model predictive control is not the only advanced process control strategy to exist. In this thesis, economic model predictive control was particularly appealing as it focuses solely on optimising the endpoint values. Yet there is certainly scope in testing the performance of other control methods including reinforcement learning which has demonstrated success in other fed-batch bioprocesses.

6.3 Concluding Remarks

Overall, the results presented in this thesis have contributed to the modelling and control of monoclonal antibody production processes. The computational work combined with experimental testing has also allowed for a realistic representation of the challenges that are present during upstream biopharmaceutical processing. Finally, the workflows proposed could enable reduced process development timelines, whilst achieving high performance, robust process, and potentially lowering the cost of process development or event manufacturing. Achieving this will create more sustainable processes, with lower environmental impacts whilst even increasing accessibility of biotherapeutics due to reduced costs.

Chapter 7 References

- Abu-Absi, N. R., Kenty, B. M., Cuellar, M. E., Borys, M. C., Sakhamuri, S., Strachan, D. J., Hausladen, M. C., & Li, Z. J. (2011). Real-time monitoring of multiple parameters in mammalian cell culture bioreactors using an in-line Raman spectroscopy probe. *Biotechnology and Bioengineering*, 108(5), 1215–1221. <https://doi.org/10.1002/bit.23023>
- Aehle, M., Bork, K., Schaepe, S., Kuprijanov, A., Horstkorte, R., Simutis, R., & Lübbert, A. (2012). Increasing batch-to-batch reproducibility of CHO-cell cultures using a model predictive control approach. *Cytotechnology*, 64(6), 623–634. <https://doi.org/10.1007/s10616-012-9438-1>
- Aehle, M., Schaepe, S., Kuprijanov, A., Simutis, R., & Lübbert, A. (2011). Simple and efficient control of CHO cell cultures. *Journal of Biotechnology*, 153(1–2), 56–61. <https://doi.org/10.1016/j.jbiotec.2011.03.006>
- Affairs, U. Nations. D. of E. and S. (2022). *The Sustainable Development Goals: Report 2022*. UN.
- Ahleboot, Z., Khorshidtalab, M., Motahari, P., Mahboudi, R., Arjmand, R., Mokarizadeh, A., & Maleknia, S. (2021). Designing a strategy for ph control to improve cho cell productivity in bioreactor. *Avicenna Journal of Medical Biotechnology*, 13(3), 123–130. <https://doi.org/10.18502/ajmb.v13i3.6365>
- Alhuthali, S., & Kontoravdi, C. (2022). Population balance modelling captures host cell protein dynamics in CHO cell cultures. *PLoS ONE*, 17(3 March), 1–25. <https://doi.org/10.1371/journal.pone.0265886>

- Alhuthali, S., Kotidis, P., & Kontoravdi, C. (2021). Osmolality effects on cho cell growth, cell volume, antibody productivity and glycosylation. *International Journal of Molecular Sciences*, 22(7). <https://doi.org/10.3390/ijms22073290>
- Altamirano, C., Berrios, J., Vergara, M., & Becerra, S. (2013). Advances in improving mammalian cells metabolism for recombinant protein production. *Electronic Journal of Biotechnology*, 16(3). <https://doi.org/10.2225/vol16-issue3-fulltext-2>
- Altamirano, C., Paredes, C., Cairó, J. J., & Gòdia, F. (2000). Improvement of CHO cell culture medium formulation: Simultaneous substitution of glucose and glutamine. *Biotechnology Progress*, 16(1), 69–75. <https://doi.org/10.1021/bp990124j>
- Amgen. (2023). *AMJEVITA™ (ADALIMUMAB-ATTO), FIRST BIOSIMILAR TO HUMIRA®, NOW AVAILABLE IN THE UNITED STATES*. <https://www.amgen.com/newsroom/press-releases/2023/01/amjevita-adalimumabatto-first-biosimilar-to-humira-now-available-in-the-united-states>
- Anane, E., Knudsen, I. M., & Wilson, G. C. (2021). Scale-down cultivation in mammalian cell bioreactors—The effect of bioreactor mixing time on the response of CHO cells to dissolved oxygen gradients. *Biochemical Engineering Journal*, 166(December 2020), 1–8. <https://doi.org/10.1016/j.bej.2020.107870>
- Anurag S Rathore, & Helen Winkle. (2009). Quality-by-design approach. *Nature Biotechnology*, 27(1), 26–34. <https://doi.org/10.1038/nbt0109-26>
- Ball, P. (2021). The lightning-fast quest for COVID vaccines - and what it means for other diseases. *Nature*, 589(7840), 16–18. <https://doi.org/10.1038/d41586-020-03626-1>
- Barbosa, R. (2021). *Development of hybrid platforms for antibody manufacturing optimisation* (Issue March). Imperial College London.
- Bauer, M., & Craig, I. K. (2008a). Economic assessment of advanced process control - A survey and framework. *Journal of Process Control*, 18(1), 2–18. <https://doi.org/10.1016/j.jprocont.2007.05.007>

- Bauer, M., & Craig, I. K. (2008b). Economic assessment of advanced process control - A survey and framework. *Journal of Process Control*, 18(1), 2–18. <https://doi.org/10.1016/j.jprocont.2007.05.007>
- Bayer, B., Duerkop, M., Pörtner, R., & Möller, J. (2023). Comparison of mechanistic and hybrid modeling approaches for characterization of a CHO cultivation process: Requirements, pitfalls and solution paths. *Biotechnology Journal*, 18(1), 1–12. <https://doi.org/10.1002/biot.202200381>
- Berry, B. N., Dobrowsky, T. M., Timson, R. C., Kshirsagar, R., Ryll, T., & Wiltberger, K. (2016). Quick generation of Raman spectroscopy based in-process glucose control to influence biopharmaceutical protein product quality during mammalian cell culture. *Biotechnology Progress*, 32(1), 224–234. <https://doi.org/10.1002/btpr.2205>
- Bibila, T. A., & Robinson, D. K. (1995). In Pursuit of the Optimal Fed-Batch Process for Monoclonal Antibody Production. *Biotechnology Progress*, 11(1), 1–13. <https://doi.org/10.1021/bp00031a001>
- Biechele, P., Busse, C., Solle, D., Scheper, T., & Reardon, K. (2015). Sensor systems for bioprocess monitoring. *Engineering in Life Sciences*, 15(5), 469–488. <https://doi.org/10.1002/elsc.201500014>
- Bolmanis, E., Dubencovs, K., Suleiko, A., & Vanags, J. (2023). Model Predictive Control—A Stand Out among Competitors for Fed-Batch Fermentation Improvement. *Fermentation*, 9(3), 1–28. <https://doi.org/10.3390/fermentation9030206>
- Buchsteiner, M., Quek, L. E., Gray, P., & Nielsen, L. K. (2018). Improving culture performance and antibody production in CHO cell culture processes by reducing the Warburg effect. *Biotechnology and Bioengineering*, 115(9), 2315–2327. <https://doi.org/10.1002/bit.26724>
- Buss, N. A. P. S., Henderson, S. J., McFarlane, M., Shenton, J. M., & De Haan, L. (2012). Monoclonal antibody therapeutics: History and future. *Current Opinion in Pharmacology*, 12(5), 615–622. <https://doi.org/10.1016/j.coph.2012.08.001>

- Bydlinski, N., Maresch, D., Schmieder, V., Klanert, G., Strasser, R., & Borth, N. (2018). The contributions of individual galactosyltransferases to protein specific N-glycan processing in Chinese Hamster Ovary cells. *Journal of Biotechnology*, 282(July), 101–110. <https://doi.org/10.1016/j.jbiotec.2018.07.015>
- Camacho, E. F., Bordons, C., Camacho, E. F., & Bordons, C. (2007). Nonlinear model predictive control. *Model Predictive Control*, 249–288.
- Chames, P., Van Regenmortel, M., Weiss, E., & Baty, D. (2009). Therapeutic antibodies: Successes, limitations and hopes for the future. *British Journal of Pharmacology*, 157(2), 220–233. <https://doi.org/10.1111/j.1476-5381.2009.00190.x>
- Chan, F. K.-M., Moriwaki, K., & Rosa, M. J. De. (2013). Detection of Necrosis by Release of Lactate Dehydrogenase Activity. In *Methods in Molecular Biology*.
- Chen, G., Hu, J., Qin, Y., & Zhou, W. (2021). Viable cell density on-line auto-control in perfusion cell culture aided by in-situ Raman spectroscopy. *Biochemical Engineering Journal*, 172(December 2020), 108063. <https://doi.org/10.1016/j.bej.2021.108063>
- Chen, N., Bennett, M. H., & Kontoravdi, C. (2014). Analysis of Chinese hamster ovary cell metabolism through a combined computational and experimental approach. *Cytotechnology*, 66(6), 945–966. <https://doi.org/10.1007/s10616-013-9648-1>
- Chopda, V., Gyorgypal, A., Yang, O., Singh, R., Ramachandran, R., Zhang, H., Tsilomelekis, G., Chundawat, S. P. S., & Ierapetritou, M. G. (2022). Recent advances in integrated process analytical techniques, modeling, and control strategies to enable continuous biomanufacturing of monoclonal antibodies. *Journal of Chemical Technology and Biotechnology*, 97(9), 2317–2335. <https://doi.org/10.1002/jctb.6765>
- Chotteau, V., & Hjalmarsson, H. (2011). Tuning of dissolved oxygen and pH PID control parameters in large scale bioreactor by lag control. *Proceedings of the 21st Annual Meeting of the European Society for Animal Cell Technology (ESACT), Dublin, Ireland, June 7-10, 2009*, 327–330.
- Claßen, J., Aupert, F., Reardon, K. F., Solle, D., & Scheper, T. (2017). Spectroscopic sensors for in-line bioprocess monitoring in research and pharmaceutical industrial application.

- Analytical and Bioanalytical Chemistry*, 409(3), 651–666.
<https://doi.org/10.1007/s00216-016-0068-x>
- Coulet, M., Kepp, O., Kroemer, G., & Basmaciogullari, S. (2022). Metabolic Profiling of CHO Cells during the Production of Biotherapeutics. *Cells*, 11(12), 1–21.
<https://doi.org/10.3390/cells11121929>
- Craven, S., Shirsat, N., Whelan, J., & Glennon, B. (2013). Process model comparison and transferability across bioreactor scales and modes of operation for a mammalian cell bioprocess. *Biotechnology Progress*, 29(1), 186–196. <https://doi.org/10.1002/btpr.1664>
- Craven, S., Whelan, J., & Glennon, B. (2014). Glucose concentration control of a fed-batch mammalian cell bioprocess using a nonlinear model predictive controller. *Journal of Process Control*, 24(4), 344–357. <https://doi.org/10.1016/j.jprocont.2014.02.007>
- Das, N. (2016). Commercializing high-concentration mAbs. *BioPharm International*, 29(11), 47–49.
- Dewasme, L., Mäkinen, M., & Chotteau, V. (2023). Practical data-driven modeling and robust predictive control of mammalian cell fed-batch process. *Computers and Chemical Engineering*, 171(July 2022), 108164.
<https://doi.org/10.1016/j.compchemeng.2023.108164>
- Domján, J., Fricška, A., Madarász, L., Gyürkés, M., Köte, Á., Farkas, A., Vass, P., Fehér, C., Horváth, B., Könczöl, K., Pataki, H., Nagy, Z. K., Marosi, G. J., & Hirsch, E. (2020). Raman-based dynamic feeding strategies using real-time glucose concentration monitoring system during adalimumab producing CHO cell cultivation. *Biotechnology Progress*, 36(6), 1–12. <https://doi.org/10.1002/btpr.3052>
- Domján, J., Pantea, E., Gyürkés, M., Madarász, L., Kozák, D., Farkas, A., Horváth, B., Benkő, Z., Nagy, Z. K., Marosi, G., & Hirsch, E. (2022). Real-time amino acid and glucose monitoring system for the automatic control of nutrient feeding in CHO cell culture using Raman spectroscopy. *Biotechnology Journal*, 17(5), 1–14.
<https://doi.org/10.1002/biot.202100395>

- Dong, K., & Ma, T. (2022). *First Steps Toward Understanding the Extrapolation of Nonlinear Models to Unseen Domains*. <http://arxiv.org/abs/2211.11719>
- Downey, B., Schmitt, J., Beller, J., Russell, B., Quach, A., Hermann, E., Lyon, D., & Breit, J. (2017). Process-aware data-driven modelling and model predictive control of bioreactor for the production of monoclonal antibodies. *Biotechnology Progress*, 33(6), 1647–1661. <https://doi.org/10.1002/btpr.2537>
- Doyle, K., Tsopanoglou, A., Fejér, A., Glennon, B., & del Val, I. J. (2023). Automated assembly of hybrid dynamic models for CHO cell culture processes. *Biochemical Engineering Journal*, 191(July 2022). <https://doi.org/10.1016/j.bej.2022.108763>
- Ellis, M., Durand, H., & Christofides, P. D. (2014). A tutorial review of economic model predictive control methods. *Journal of Process Control*, 24(8), 1156–1178. <https://doi.org/10.1016/j.jprocont.2014.03.010>
- Eyster, T. W., Talwar, S., Fernandez, J., Foster, S., Hayes, J., Allen, R., Reidinger, S., Wan, B., Ji, X., Aon, J., Patel, P., & Ritz, D. B. (2021). Tuning monoclonal antibody galactosylation using Raman spectroscopy-controlled lactic acid feeding. *Biotechnology Progress*, 37(1), 1–13. <https://doi.org/10.1002/btpr.3085>
- Fan, Y., Jimenez Del Val, I., Müller, C., Wagtberg Sen, J., Rasmussen, S. K., Kontoravdi, C., Weilguny, D., & Andersen, M. R. (2015). Amino acid and glucose metabolism in fed-batch CHO cell culture affects antibody production and glycosylation. *Biotechnology and Bioengineering*, 112(3), 521–535. <https://doi.org/10.1002/bit.25450>
- FDA. (2004). *Guidance for Industry, PAT-A Framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance*. September. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070305.pdf>
- Fisher, O. J., Watson, N. J., Porcu, L., Bacon, D., Rigley, M., & Gomes, R. L. (2022). Data-driven modelling for resource recovery: Data volume, variability, and visualisation for an industrial bioprocess. *Biochemical Engineering Journal*, 185(January), 108499. <https://doi.org/10.1016/j.bej.2022.108499>

- Fouladiha, H., Marashi, S. A., Torkashvand, F., Mahboudi, F., Lewis, N. E., & Vaziri, B. (2020). A metabolic network-based approach for developing feeding strategies for CHO cells to increase monoclonal antibody production. *Bioprocess and Biosystems Engineering*, 43(8), 1381–1389. <https://doi.org/10.1007/s00449-020-02332-6>
- Fox, S. R., Patel, U. A., Yap, M. G. S., & Wang, D. I. C. (2004). Maximizing Interferon- γ Production by Chinese Hamster Ovary Cells Through Temperature Shift Optimization: Experimental and Modeling. *Biotechnology and Bioengineering*, 85(2), 177–184. <https://doi.org/10.1002/bit.10861>
- Fung Shek, C., Kotidis, P., & Betenbaugh, M. (2021). Mechanistic and data-driven modeling of protein glycosylation. *Current Opinion in Chemical Engineering*, 32, 100690. <https://doi.org/10.1016/j.coche.2021.100690>
- Gadgil, M. (2015). Development of a mathematical model for animal cell culture without pH control and its application for evaluation of clone screening outcomes in shake flask culture. *Journal of Chemical Technology and Biotechnology*, 90(1), 166–175. <https://doi.org/10.1002/jctb.4302>
- Gagnon, M., Hiller, G., Luan, Y. T., Kittredge, A., Defelice, J., & Drapeau, D. (2011). High-End pH-controlled delivery of glucose effectively suppresses lactate accumulation in CHO Fed-batch cultures. *Biotechnology and Bioengineering*, 108(6), 1328–1337. <https://doi.org/10.1002/bit.23072>
- Garrido, J., Vázquez, F., & Morilla, F. (2016). Multivariable PID control by decoupling. *International Journal of Systems Science*, 47(5), 1054–1072. <https://doi.org/10.1080/00207721.2014.911390>
- Gibbons, L., Maslanka, F., Le, N., Magill, A., Singh, P., Mclaughlin, J., Madden, F., Hayes, R., McCarthy, B., Rode, C., O'Mahony, J., Rea, R., & O'Mahony-Hartnett, C. (2023). An assessment of the impact of Raman based glucose feedback control on CHO cell bioreactor process development. *Biotechnology Progress*, 39(5), 1–14. <https://doi.org/10.1002/btpr.3371>

- Glass, H. E., Kolassa, E. M., & Muniz, E. (2016). Drug development through modeling and simulation-The business case. *Applied Clinical Trials*, 25.
- Glassy, M. C., & Gupta, R. (2014). Technical and ethical limitations in making human monoclonal antibodies (an overview). *Human Monoclonal Antibodies: Methods and Protocols*, 9–36.
- Goldrick, S., Lee, K., Spencer, C., Holmes, W., Kuiper, M., Turner, R., & Farid, S. S. (2018). On-Line Control of Glucose Concentration in High-Yielding Mammalian Cell Cultures Enabled Through Oxygen Transfer Rate Measurements. *Biotechnology Journal*, 13(4). <https://doi.org/10.1002/biot.201700607>
- Graf, A., Lemke, J., Schulze, M., Soeldner, R., Rebner, K., Hoehse, M., & Matuszczyk, J. (2022). A Novel Approach for Non-Invasive Continuous In-Line Control of Perfusion Cell Cultivations by Raman Spectroscopy. *Frontiers in Bioengineering and Biotechnology*, 10(April), 1–11. <https://doi.org/10.3389/fbioe.2022.719614>
- Gros, S., Zanon, M., Quirynen, R., Bemporad, A., & Diehl, M. (2020). From linear to nonlinear MPC: bridging the gap via the real-time iteration. *International Journal of Control*, 93(1), 62–80. <https://doi.org/10.1080/00207179.2016.1222553>
- Guerra, A., von Stosch, M., & Glassey, J. (2019). Toward biotherapeutic product real-time quality monitoring. *Critical Reviews in Biotechnology*, 39(3), 289–305. <https://doi.org/10.1080/07388551.2018.1524362>
- Ha, T. K., Kim, D., Kim, C. L., Grav, L. M., & Lee, G. M. (2022). Factors affecting the quality of therapeutic proteins in recombinant Chinese hamster ovary cell culture. *Biotechnology Advances*, 54(June 2021), 107831. <https://doi.org/10.1016/j.biotechadv.2021.107831>
- Harcum, S. W., Elliott, K. S., Skelton, B. A., Klaubert, S. R., Dahodwala, H., & Lee, K. H. (2022). PID controls: the forgotten bioprocess parameters. *Discover Chemical Engineering*, 2(1). <https://doi.org/10.1007/s43938-022-00008-z>
- Hartley, F., Walker, T., Chung, V., & Morten, K. (2018). Mechanisms driving the lactate switch in Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 115(8), 1890–1903. <https://doi.org/10.1002/bit.26603>

- Hebing, L., Tran, F., Brandt, H., & Engell, S. (2020). Robust Optimizing Control of Fermentation Processes Based on a Set of Structurally Different Process Models. *Industrial and Engineering Chemistry Research*, 59(6), 2566–2580. <https://doi.org/10.1021/acs.iecr.9b05504>
- Helgers, H., Schmidt, A., & Strube, J. (2022). Towards Autonomous Process Control—Digital Twin for CHO Cell-Based Antibody Manufacturing Using a Dynamic Metabolic Model. *Processes*, 10(2). <https://doi.org/10.3390/pr10020316>
- Hernández Rodríguez, T., Posch, C., Schmutzhard, J., Stettner, J., Weihs, C., Pörtner, R., & Frahm, B. (2019). Predicting industrial-scale cell culture seed trains—A Bayesian framework for model fitting and parameter estimation, dealing with uncertainty in measurements and model parameters, applied to a nonlinear kinetic cell culture model, using an MCMC method. *Biotechnology and Bioengineering*, 116(11), 2944–2959. <https://doi.org/10.1002/bit.27125>
- Herwig, C. (2018). Hybrid Modelling and Multi- Parametric Control of Bioprocesses. In *Hybrid Modelling and Multi- Parametric Control of Bioprocesses*. <https://doi.org/10.3390/books978-3-03842-746-9>
- Hippach, M. B., Schwartz, I., Pei, J., Huynh, J., Kawai, Y., & Zhu, M. M. (2018). Fluctuations in dissolved oxygen concentration during a CHO cell culture process affects monoclonal antibody productivity and the sulfhydryl-drug conjugation process. *Biotechnology Progress*, 34(6), 1427–1437. <https://doi.org/10.1002/btpr.2697>
- Hogiri, T., Tamashima, H., Nishizawa, A., & Okamoto, M. (2018). Optimization of a pH-shift control strategy for producing monoclonal antibodies in Chinese hamster ovary cell cultures using a pH-dependent dynamic model. *Journal of Bioscience and Bioengineering*, 125(2), 245–250. <https://doi.org/10.1016/j.jbiosc.2017.08.015>
- Holm, P., Allesø, M., Bryder, M. C., & Holm, R. (2017). Q8 (R2) Pharmaceutical Development. *ICH Quality Guidelines: An Implementation Guide*, 535–577.

- Hoshan, L., Jiang, R., Moroney, J., Bui, A., Zhang, X., Hang, T. C., & Xu, S. (2019). Effective bioreactor pH control using only sparging gases. *Biotechnology Progress*, 35(1), 1–7. <https://doi.org/10.1002/btpr.2743>
- Hu, Q., Jiang, B., Liu, D., Tang, X., Daly, T., & Shameem, M. (2020). Chapter 14: Practical considerations in high concentration formulation development for monoclonal antibody drug products. *Development of Biopharmaceutical Drug-Device Products*, 343–372.
- Huang, Y. M., Hu, W. W., Rustandi, E., Chang, K., Yusuf-Makagiansar, H., & Ryll, T. (2010). Maximizing productivity of CHO cell-based fed-batch culture using chemically defined media conditions and typical manufacturing equipment. *Biotechnology Progress*, 26(5), 1400–1410. <https://doi.org/10.1002/btpr.436>
- Huang, Z., Xu, J., Yongky, A., Morris, C. S., Polanco, A. L., Reily, M., Borys, M. C., Li, Z. J., & Yoon, S. (2020). CHO cell productivity improvement by genome-scale modeling and pathway analysis: Application to feed supplements. *Biochemical Engineering Journal*, 160(April), 107638. <https://doi.org/10.1016/j.bej.2020.107638>
- Jarboe, J., Gupta, A., & Saif, M. W. (2014). Therapeutic human monoclonal antibodies against cancer. *Human Monoclonal Antibodies: Methods and Protocols*, 61–77.
- Jayapal, K. P., Wlaschin, K. F., Hu, W. S., & Yap, M. G. S. (2007). Recombinant protein therapeutics from CHO Cells - 20 years and counting. *Chemical Engineering Progress*, 103(10), 40–47.
- Jiang, R., Chen, H., & Xu, S. (2018). pH excursions impact CHO cell culture performance and antibody N-linked glycosylation. *Bioprocess and Biosystems Engineering*, 41(12), 1731–1741. <https://doi.org/10.1007/s00449-018-1996-y>
- Jimenez del Val, I., Fan, Y., & Weilguny, D. (2016). Dynamics of immature mAb glycoform secretion during CHO cell culture: An integrated modelling framework. *Biotechnology Journal*, 11(5), 610–623. <https://doi.org/10.1002/biot.201400663>
- Jiménez del Val, I., Kyriakopoulos, S., Albrecht, S., Stockmann, H., Rudd, P. M., Polizzi, K. M., & Kontoravdi, C. (2023). CHOmpact: A reduced metabolic model of Chinese hamster

- ovary cells with enhanced interpretability. *Biotechnology and Bioengineering*, 120(9), 2479–2493. <https://doi.org/10.1002/bit.28459>
- Jones, J., Kindembe, D., Branton, H., Lawal, N., Montero, E. L., Mack, J., Shi, S., Patton, R., & Montague, G. (2023). Improved control strategies for the environment within cell culture bioreactors. *Food and Bioprocess Processing*, 138, 209–220. <https://doi.org/10.1016/j.fbp.2023.02.004>
- Joynes, L., & Zhang, J. (2022). A Feedback Control Strategy for a Fed-Batch Monoclonal Antibody Production Process Utilising Infrequent and Irregular Sampled Measurements. *Processes*, 10(8). <https://doi.org/10.3390/pr10081448>
- Kappatou, C. D., Ehsani, A., Niedenführ, S., Mhamdi, A., Schuppert, A., & Mitsos, A. (2020). Quality-targeting dynamic optimization of monoclonal antibody production. *Computers and Chemical Engineering*, 142, 107004. <https://doi.org/10.1016/j.compchemeng.2020.107004>
- Karra, S., Sager, B., & Karim, M. N. (2010). Multi-Scale Modeling of Heterogeneities in Mammalian Cell Culture Processes. *Industrial & Engineering Chemistry Research*, 49(17), 7990–8006. <https://doi.org/10.1021/ie100125a>
- Kelley, B. (2020). Developing therapeutic monoclonal antibodies at pandemic pace. *Nature Biotechnology*, 38(5), 540–545. <https://doi.org/10.1038/s41587-020-0512-5>
- Kesik-Brodacka, M. (2018). Progress in biopharmaceutical development. *Biotechnology and Applied Biochemistry*, 65(3), 306–322. <https://doi.org/10.1002/bab.1617>
- Kinch, M. S., Kraft, Z., & Schwartz, T. (2023). Monoclonal antibodies: Trends in therapeutic success and commercial focus. *Drug Discovery Today*, 28(1), 103415. <https://doi.org/10.1016/j.drudis.2022.103415>
- Kiparissides, A., Koutinas, M., Kontoravdi, C., Mantalaris, A., & Pistikopoulos, E. N. (2011). “Closing the loop” in biological systems modeling-from the in silico to the in vitro. *Automatica*, 47(6), 1147–1155. <https://doi.org/10.1016/j.automatica.2011.01.013>

- Kiparissides, A., Kucherenko, S. S., Mantalaris, A., & Pistikopoulos, E. N. (2009). Global sensitivity analysis challenges in biological systems modeling. *Industrial and Engineering Chemistry Research*, 48(15), 7168–7180. <https://doi.org/10.1021/ie900139x>
- Kiparissides, A., Pistikopoulos, E. N., & Mantalaris, A. (2015). On the model-based optimization of secreting mammalian cell (GS-NS0) cultures. *Biotechnology and Bioengineering*, 112(3), 536–548. <https://doi.org/10.1002/bit.25457>
- Klein, T., Heinzl, N., Kroll, P., Brunner, M., Herwig, C., & Neutsch, L. (2015). Quantification of cell lysis during CHO bioprocesses: Impact on cell count, growth kinetics and productivity. *Journal of Biotechnology*, 207, 67–76. <https://doi.org/10.1016/j.jbiotec.2015.04.021>
- Köhler, G., Milstein, C. (1975). Cultivos continuos de células fusionadas que secretan anticuerpos de especificidad predefinida. *Nature*, 256(Mopc 21), 495–497. <http://www.nature.com/articles/256495a0%0Apapers3://publication/doi/10.1038/256495a0>
- Konakovsky, V., Clemens, C., Müller, M. M., Bechmann, J., Berger, M., Schlatter, S., & Herwig, C. (2016). Metabolic control in mammalian fed-batch cell cultures for reduced lactic acid accumulation and improved process robustness. *Bioengineering*, 3(1). <https://doi.org/10.3390/bioengineering3010005>
- Kontoravdi, C., Asprey, S. P., Pistikopoulos, E. N., & Mantalaris, A. (2005). Application of global sensitivity analysis to determine goals for design of experiments: An example study on antibody-producing cell cultures. *Biotechnology Progress*, 21(4), 1128–1135. <https://doi.org/10.1021/bp050028k>
- Kontoravdi, C., Pistikopoulos, E. N., & Mantalaris, A. (2010). Systematic development of predictive mathematical models for animal cell cultures. *Computers and Chemical Engineering*, 34(8), 1192–1198. <https://doi.org/10.1016/j.compchemeng.2010.03.012>
- Kopanos, G. M., & Pistikopoulos, E. N. (2014). Reactive scheduling by a multiparametric programming rolling horizon framework: A case of a network of combined heat and

- power units. *Industrial and Engineering Chemistry Research*, 53(11), 4366–4386. <https://doi.org/10.1021/ie402393s>
- Kotidis, P., Demis, P., Goey, C. H., Correa, E., McIntosh, C., Trepekli, S., Shah, N., Klymenko, O. V., & Kontoravdi, C. (2019). Constrained global sensitivity analysis for bioprocess design space identification. *Computers and Chemical Engineering*, 125, 558–568. <https://doi.org/10.1016/j.compchemeng.2019.01.022>
- Kotidis, P., Jedrzejewski, P., Sou, S. N., Sellick, C., Polizzi, K., del Val, I. J., & Kontoravdi, C. (2019). Model-based optimization of antibody galactosylation in CHO cell culture. *Biotechnology and Bioengineering*, 116(7), 1612–1626. <https://doi.org/10.1002/bit.26960>
- Kotidis, P., & Kontoravdi, C. (2020). Harnessing the potential of artificial neural networks for predicting protein glycosylation. *Metabolic Engineering Communications*, 10, e00131. <https://doi.org/10.1016/j.mec.2020.e00131>
- Kotidis, P., Pappas, I., Avraamidou, S., Pistikopoulos, E. N., Kontoravdi, C., & Papathanasiou, M. M. (2021). DigiGlyc: A hybrid tool for reactive scheduling in cell culture systems. *Computers and Chemical Engineering*, 154, 107460. <https://doi.org/10.1016/j.compchemeng.2021.107460>
- Koumpouras, G., & Kontoravdi, C. (2012). Dynamic Optimization of Bioprocesses. *Applied Mathematics*, 03(10), 1487–1495. <https://doi.org/10.4236/am.2012.330208>
- Krammer, F. (2020). SARS-CoV-2 vaccines in development. *Nature*, 586(7830), 516–527. <https://doi.org/10.1038/s41586-020-2798-3>
- Kreuter, J. Y.-T. K. M.-J. C. (2006). The effect of dissolved oxygen on the production and the glycosylation profile of recombinant human erythropoietin produced from CHO cells. *Biotechnology and Bioengineering*, 94(3), 481–494. <https://doi.org/10.1002/bit.20875>
- Kucherenko, S., & Zaccueus, O. (2021). *SobolGSA Software*.
- Kumar, D., Gangwar, N., Rathore, A. S., & Ramteke, M. (2022). Multi-objective optimization of monoclonal antibody production in bioreactor. *Chemical Engineering and Processing*

- *Process Intensification*, 180(July 2021), 108720.
<https://doi.org/10.1016/j.cep.2021.108720>
- Kunkel, J. P., Jan, D. C. H., Jamieson, J. C., & Butler, M. (1998). Dissolved oxygen concentration in serum-free continuous culture affects N-linked glycosylation of a monoclonal antibody. *Journal of Biotechnology*, 62(1), 55–71.
[https://doi.org/10.1016/S0168-1656\(98\)00044-3](https://doi.org/10.1016/S0168-1656(98)00044-3)
- Kyriakopoulos, S., Ang, K. S., Lakshmanan, M., Huang, Z., Yoon, S., Gunawan, R., & Lee, D. Y. (2018). Kinetic Modeling of Mammalian Cell Culture Bioprocessing: The Quest to Advance Biomanufacturing. *Biotechnology Journal*, 13(3).
<https://doi.org/10.1002/biot.201700229>
- Kyriakopoulos, S., & Kontoravdi, C. (2014). A framework for the systematic design of fed-batch strategies in mammalian cell culture. *Biotechnology and Bioengineering*, 111(12), 2466–2476. <https://doi.org/10.1002/bit.25319>
- Lalonde, M. E., & Durocher, Y. (2017). Therapeutic glycoprotein production in mammalian cells. *Journal of Biotechnology*, 251(April), 128–140.
<https://doi.org/10.1016/j.jbiotec.2017.04.028>
- Lam, C. M., Sriyudthsak, K., Kontoravdi, C., Kothari, K., Park, H. H., Pistikopoulos, E. N., & Mantalaris, A. (2008). Cell cycle modelling for offline dynamic optimisation of mammalian cultures. *Computer Aided Chemical Engineering*, 25, 109–114.
[https://doi.org/10.1016/S1570-7946\(08\)80023-5](https://doi.org/10.1016/S1570-7946(08)80023-5)
- Langheinrich, C., & Nienow, A. W. (1999). Control of pH in large-scale, free suspension animal cell bioreactors: Alkali addition and pH excursions. *Biotechnology and Bioengineering*, 66(3), 171–179. [https://doi.org/10.1002/\(SICI\)1097-0290\(1999\)66:3<171::AID-BIT5>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1097-0290(1999)66:3<171::AID-BIT5>3.0.CO;2-T)
- Lao, M. S., & Toth, D. (1997). Effects of ammonium and lactate on growth and metabolism of a recombinant Chinese hamster ovary cell culture. *Biotechnology Progress*, 13(5), 688–691. <https://doi.org/10.1021/bp9602360>

- Le Gal, K., Schmidt, E. E., & Sayin, V. I. (2021). Cellular redox homeostasis. *Antioxidants*, 10(9), 1–7. <https://doi.org/10.3390/antiox10091377>
- Le, H., Kabbur, S., Pollastrini, L., Sun, Z., Mills, K., Johnson, K., Karypis, G., & Hu, W. S. (2012). Multivariate analysis of cell culture bioprocess data—Lactate consumption as process indicator. *Journal of Biotechnology*, 162(2–3), 210–223. <https://doi.org/10.1016/j.jbiotec.2012.08.021>
- Lee, A. P., Kok, Y. J., Lakshmanan, M., Leong, D., Zheng, L., Lim, H. L., Chen, S., Mak, S. Y., Ang, K. S., Templeton, N., Salim, T., Wei, X., Gifford, E., Tan, A. H. M., Bi, X., Ng, S. K., Lee, D. Y., Ling, W. L. W., & Ho, Y. S. (2021). Multi-omics profiling of a CHO cell culture system unravels the effect of culture pH on cell growth, antibody titer, and product quality. *Biotechnology and Bioengineering*, 118(11), 4305–4316. <https://doi.org/10.1002/bit.27899>
- Li, B. T., Smit, E. F., Goto, Y., Nakagawa, K., Udagawa, H., Mazières, J., Nagasaka, M., Bazhenova, L., Saltos, A. N., Felip, E., Pacheco, J. M., Pérol, M., Paz-Ares, L., Saxena, K., Shiga, R., Cheng, Y., Acharyya, S., Vitazka, P., Shahidi, J., ... Jänne, P. A. (2022). Trastuzumab Deruxtecan in HER2 -Mutant Non–Small-Cell Lung Cancer . *New England Journal of Medicine*, 386(3), 241–251. <https://doi.org/10.1056/nejmoa2112431>
- Li, F., Vijayasankaran, N., Shen, A., Kiss, R., & Amanullah, A. (2010). Cell culture processes for monoclonal antibody production. *MAbs*, 2(5), 466–479. <https://doi.org/10.4161/mabs.2.5.12720>
- Li, J., Wong, C. L., Vijayasankaran, N., Hudson, T., & Amanullah, A. (2012). Feeding lactate for CHO cell culture processes: Impact on culture metabolism and performance. *Biotechnology and Bioengineering*, 109(5), 1173–1186. <https://doi.org/10.1002/bit.24389>
- Li, M., Ebel, B., Chauchard, F., Guédon, E., & Marc, A. (2018). Parallel comparison of in situ Raman and NIR spectroscopies to simultaneously measure multiple variables toward real-time monitoring of CHO cell bioreactor cultures. *Biochemical Engineering Journal*, 137, 205–213. <https://doi.org/10.1016/j.bej.2018.06.005>

- Liu, L. (2015). Antibody glycosylation and its impact on the pharmacokinetics and pharmacodynamics of monoclonal antibodies and Fc-fusion proteins. *Journal of Pharmaceutical Sciences*, 104(6), 1866–1884. <https://doi.org/10.1002/jps.24444>
- Logist, F., Houska, B., Diehl, M., & Van Impe, J. F. (2010). A toolkit for multi-objective optimal control in bioprocess engineering. In *IFAC Proceedings Volumes (IFAC-PapersOnline)* (Vol. 11, Issue PART 1). IFAC. <https://doi.org/10.3182/20100707-3-BE-2012.0063>
- Looby, M., Ibarra, N., Pierce, J. J., Buckley, K., O'Donovan, E., Heenan, M., Moran, E., Farid, S. S., & Baganz, F. (2011). Application of quality by design principles to the development and technology transfer of a major process improvement for the manufacture of a recombinant protein. *Biotechnology Progress*, 27(6), 1718–1729. <https://doi.org/10.1002/btpr.672>
- Lu, F., Toh, P. C., Burnett, I., Li, F., Hudson, T., Amanullah, A., & Li, J. (2013). Automated dynamic fed-batch process and media optimization for high productivity cell culture process development. *Biotechnology and Bioengineering*, 110(1), 191–205. <https://doi.org/10.1002/bit.24602>
- Luginsland, M., Kontoravdi, C., Racher, A., Jaques, C., & Kiparissides, A. (2024). Elucidating lactate metabolism in industrial CHO cultures through the combined use of flux balance and principal component analyses. *Biochemical Engineering Journal*, 202(September 2023), 109184. <https://doi.org/10.1016/j.bej.2023.109184>
- Luo, Y., Kurian, V., & Ogunnaike, B. A. (2021). Bioprocess systems analysis, modeling, estimation, and control. *Current Opinion in Chemical Engineering*, 33, 100705. <https://doi.org/10.1016/j.coche.2021.100705>
- Luo, Y., Kurian, V., Song, L., Wells, E. A., Robinson, A. S., & Ogunnaike, B. A. (2023). Model-based control of titer and glycosylation in fed-batch mAb production: Modeling and control system development. *AIChE Journal*, 69(4), 1–17. <https://doi.org/10.1002/aic.18075>

- Luo, Y., Lovelett, R. J., Price, J. V., Radhakrishnan, D., Barnthouse, K., Hu, P., Schaefer, E., Cunningham, J., Lee, K. H., Shivappa, R. B., & Ogunnaike, B. A. (2021). Modeling the Effect of Amino Acids and Copper on Monoclonal Antibody Productivity and Glycosylation: A Modular Approach. *Biotechnology Journal*, 16(2). <https://doi.org/10.1002/biot.202000261>
- Luo, Y., Stanton, D. A., Sharp, R. C., Parrillo, A. J., Morgan, K. T., Ritz, D. B., & Talwar, S. (2023). Efficient optimization of time-varying inputs in a fed-batch cell culture process using design of dynamic experiments. *Biotechnology Progress*, May, 1–10. <https://doi.org/10.1002/btpr.3380>
- MacDonald, G. J. (2020). Biomanufacturing Makes Sense of the Industry 4.0 Concept. *Genetic Engineering and Biotechnology News*, 40(S3), S7–S10. <https://doi.org/10.1089/gen.40.S3.03>
- Maguire, J., & Peng, D. (2015). How to identify critical quality attributes and critical process parameters. *Office of Process and Facility (OPF) OPQ/CDER/FDA, FDA/PQRI 2nd Conference, North Bethesda, MD.*
- Majumdar, S., Desai, R., Hans, A., Dandekar, P., & Jain, R. (2024). From Efficiency to Yield : Exploring Recent Advances in CHO Cell Line Development for Monoclonal Antibodies. *Molecular Biotechnology*, 0123456789. <https://doi.org/10.1007/s12033-024-01060-6>
- Maria, G. (2020). Model-based optimization of a fed-batch bioreactor for mab production using a hybridoma cell culture. *Molecules*, 25(23), 1–25. <https://doi.org/10.3390/molecules25235648>
- Martens, D. E., Van Den End, E. J., & Streefland, M. (2014). Configuration of bioreactors. *Methods in Molecular Biology*, 1104, 285–311. https://doi.org/10.1007/978-1-62703-733-4_19
- Martínez, V. S., Dietmair, S., Quek, L. E., Hodson, M. P., Gray, P., & Nielsen, L. K. (2013). Flux balance analysis of CHO cells before and after a metabolic switch from lactate production to consumption. *Biotechnology and Bioengineering*, 110(2), 660–666. <https://doi.org/10.1002/bit.24728>

- Martínez-Monge, I., Comas, P., Triquell, J., Casablancas, A., Lecina, M., Paredes, C. J., & Cairó, J. J. (2019). Concomitant consumption of glucose and lactate: A novel batch production process for CHO cells. *Biochemical Engineering Journal*, 151(April), 107358. <https://doi.org/10.1016/j.bej.2019.107358>
- Martínez-Monge, I., Martínez, C., Decker, M., Udugama, I. A., Marín de Mas, I., Gernaey, K. V., & Nielsen, L. K. (2022). Soft-sensors application for automated feeding control in high-throughput mammalian cell cultures. *Biotechnology and Bioengineering*, 119(4), 1077–1090. <https://doi.org/10.1002/bit.28032>
- Masana, M. O., & Baranyi, J. (2000). Adding new factors to predictive models: The effect on the risk of extrapolation. *Food Microbiology*, 17(4), 367–374. <https://doi.org/10.1006/fmic.1999.0326>
- Matthews, T. E., Berry, B. N., Smelko, J., Moretto, J., Moore, B., & Wiltberger, K. (2016). Closed loop control of lactate concentration in mammalian cell culture by Raman spectroscopy leads to improved cell density, viability, and biopharmaceutical protein production. *Biotechnology and Bioengineering*, 113(11), 2416–2424. <https://doi.org/10.1002/bit.26018>
- McHugh, K. P., Xu, J., Aron, K. L., Borys, M. C., & Li, Z. J. (2020). Effective temperature shift strategy development and scale confirmation for simultaneous optimization of protein productivity and quality in Chinese hamster ovary cells. *Biotechnology Progress*, 36(3), 1–11. <https://doi.org/10.1002/btpr.2959>
- Mears, L., Stocks, S. M., Albaek, M. O., Sin, G., & Gernaey, K. V. (2017). Mechanistic Fermentation Models for Process Design, Monitoring, and Control. *Trends in Biotechnology*, 35(10), 914–924. <https://doi.org/10.1016/j.tibtech.2017.07.002>
- Mehdizadeh, H., Lauri, D., Karry, K. M., Moshgbar, M., Procopio-Melino, R., & Drapeau, D. (2015). Generic Raman-based calibration models enabling real-time monitoring of cell culture bioreactors. *Biotechnology Progress*, 31(4), 1004–1013. <https://doi.org/10.1002/btpr.2079>

- Mellahi, K., Brochu, D., Gilbert, M., Perrier, M., Ansorge, S., Durocher, Y., & Henry, O. (2019). Assessment of fed-batch cultivation strategies for an inducible CHO cell line. *Journal of Biotechnology*, 298(April), 45–56. <https://doi.org/10.1016/j.jbiotec.2019.04.005>
- Michl, J., Park, K. C., & Swietach, P. (2019). Evidence-based guidelines for controlling pH in mammalian live-cell culture systems. *Communications Biology*, 2(1), 1–12. <https://doi.org/10.1038/s42003-019-0393-7>
- Miller, W. M., Blanch, H. W., & Wilke, C. R. (1988). A kinetic analysis of hybridoma growth and metabolism in batch and continuous suspension culture: effect of nutrient concentration, dilution rate, and pH. *Biotechnology and Bioengineering*, 67(6).
- Mitra, S., & Murthy, G. S. (2022). Bioreactor control systems in the biopharmaceutical industry: a critical perspective. *Systems Microbiology and Biomanufacturing*, 2(1), 91–112. <https://doi.org/10.1007/s43393-021-00048-6>
- Möller, J., Kuchemüller, K. B., Steinmetz, T., Koopmann, K. S., & Pörtner, R. (2019). Model-assisted Design of Experiments as a concept for knowledge-based bioprocess development. *Bioprocess and Biosystems Engineering*, 42(5), 867–882. <https://doi.org/10.1007/s00449-019-02089-7>
- Möllerl, J., Eibl, R., Eibl, D., & Pörtner, R. (2015). Model-based DoE for feed batch cultivation of a CHO cell line. *BMC Proceedings*, 9(S9), 1–2. <https://doi.org/10.1186/1753-6561-9-s9-p42>
- Monteiro, M., Fadda, S., & Kontoravdi, C. (2023). Towards advanced bioprocess optimization: A multiscale modelling approach. *Computational and Structural Biotechnology Journal*, 21(February), 3639–3655. <https://doi.org/10.1016/j.csbj.2023.07.003>
- Moore, A., Mercer, J., Dutina, G., Donahue, C. J., Bauer, K. D., Mather, J. P., Etcheverry, T., & Ryll, T. (1997). Effects of temperature shift on cell cycle, apoptosis and nucleotide pools in CHO cell batch cultures. *Cytotechnology*, 23(1–3), 47–54. <https://doi.org/10.1023/a:1007919921991>

- Moore, B., Sanford, R., & Zhang, A. (2019). Case study: The characterization and implementation of dielectric spectroscopy (biocapacitance) for process control in a commercial GMP CHO manufacturing process. *Biotechnology Progress*, 35(3). <https://doi.org/10.1002/btpr.2782>
- Mullard, A. (2021). FDA approves 100th monoclonal antibody product. *Nature Reviews. Drug Discovery*, 20(7), 491–495. <https://doi.org/10.1038/d41573-021-00079-7>
- Naderi, S., Nikdel, A., Meshram, M., Mcconkey, B., Ingalls, B., Budman, H., & Scharer, J. (2014). Modeling of cell culture damage and recovery leads to increased antibody and biomass productivity in CHO cell cultures. *Biotechnology Journal*, 9(9), 1152–1163. <https://doi.org/10.1002/biot.201300287>
- Narayanan, H., Luna, M., Sokolov, M., Butté, A., & Morbidelli, M. (2022). Hybrid Models Based on Machine Learning and an Increasing Degree of Process Knowledge: Application to Cell Culture Processes. *Industrial and Engineering Chemistry Research*, 61(25), 8658–8672. <https://doi.org/10.1021/acs.iecr.1c04507>
- Narayanan, H., Sokolov, M., Morbidelli, M., & Butté, A. (2019). A new generation of predictive models: The added value of hybrid models for manufacturing processes of therapeutic proteins. *Biotechnology and Bioengineering*, 116(10), 2540–2549. <https://doi.org/10.1002/bit.27097>
- Narayanan, H., Sponchioni, M., & Morbidelli, M. (2022). Integration and digitalization in the manufacturing of therapeutic proteins. *Chemical Engineering Science*, 248, 117159. <https://doi.org/10.1016/j.ces.2021.117159>
- Navratil, P., Pekar, L., & Matusu, R. (2020). Control of a Multivariable System Using Optimal Control Pairs: A Quadruple-Tank Process. *IEEE Access*, 8, 2537–2563. <https://doi.org/10.1109/ACCESS.2019.2962302>
- Nikita, S., Mishra, S., Gupta, K., Runkana, V., Gomes, J., & Rathore, A. S. (2023). Advances in bioreactor control for production of biotherapeutic products. *Biotechnology and Bioengineering*, 120(5), 1189–1214. <https://doi.org/10.1002/bit.28346>

- Nimmegeers, P., Bhonsale, S., Telen, D., & Van Impe, J. (2020). Optimal experiment design under parametric uncertainty: A comparison of a sensitivities based approach versus a polynomial chaos based stochastic approach. *Chemical Engineering Science*, 221, 115651. <https://doi.org/10.1016/j.ces.2020.115651>
- Nolan, R. P., & Lee, K. (2011). Dynamic model of CHO cell metabolism. *Metabolic Engineering*, 13(1), 108–124. <https://doi.org/10.1016/j.ymben.2010.09.003>
- Oguchi, S., Saito, H., Tsukahara, M., & Tsumura, H. (2006). pH Condition in temperature shift cultivation enhances cell longevity and specific hMab productivity in CHO culture. *Cytotechnology*, 52(3), 199–207. <https://doi.org/10.1007/s10616-007-9059-2>
- Ohkubo, T., Sakumura, Y., & Kunida, K. (2023). On-Line Reoptimization of Mammalian Fed-Batch Culture Using a Nonlinear Model Predictive Controller. *New Generation Computing*, 0123456789. <https://doi.org/10.1007/s00354-023-00235-0>
- Okamura, K., Badr, S., Murakami, S., & Sugiyama, H. (2022). Hybrid Modeling of CHO Cell Cultivation in Monoclonal Antibody Production with an Impurity Generation Module. *Industrial and Engineering Chemistry Research*, 61(40), 14898–14909. <https://doi.org/10.1021/acs.iecr.2c00736>
- O'Mara, B., Gao, Z. H., Kuruganti, M., Mallett, R., Nayar, G., Smith, L., Meyer, J. D., Theriault, J., Miller, C., Cisney, J., & Fann, J. (2019). Impact of depth filtration on disulfide bond reduction during downstream processing of monoclonal antibodies from CHO cell cultures. *Biotechnology and Bioengineering*, 116(7), 1669–1683. <https://doi.org/10.1002/bit.26964>
- Ozturk, S., & Hu, W. S. (2005). Cell Culture Technology for Pharmaceutical and Cell-Based Therapies. In *Cell Culture Technology for Pharmaceutical and Cell-Based Therapies*. <https://doi.org/10.1201/9780849351068>
- Papathanasiou, M. M., & Kontoravdi, C. (2020). Engineering challenges in therapeutic protein product and process design. *Current Opinion in Chemical Engineering*, 27, 81–88. <https://doi.org/10.1016/j.coche.2019.11.010>

- Papathanasiou, M. M., Quiroga-Campano, A. L., Steinebach, F., Elviro, M., Mantalaris, A., & Pistikopoulos, E. N. (2017). Advanced model-based control strategies for the intensification of upstream and downstream processing in mAb production. *Biotechnology Progress*, 33(4), 966–988. <https://doi.org/10.1002/btpr.2483>
- Pappenreiter, M., Döbele, S., Striedner, G., Jungbauer, A., & Sissolak, B. (2022). Model predictive control for steady-state performance in integrated continuous bioprocesses. *Bioprocess and Biosystems Engineering*, 45(9), 1499–1513. <https://doi.org/10.1007/s00449-022-02759-z>
- Pereira, S., Kildegaard, H. F., & Andersen, M. R. (2018). Impact of CHO Metabolism on Cell Growth and Protein Production: An Overview of Toxic and Inhibiting Metabolites and Nutrients. *Biotechnology Journal*, 13(3), 1–13. <https://doi.org/10.1002/biot.201700499>
- Pham, T. D., Manapragada, C., Sun, Y., Bassett, R., & Aickelin, U. (2023). A scoping review of supervised learning modelling and data-driven optimisation in monoclonal antibody process development. *Digital Chemical Engineering*, 7(December 2022), 100080. <https://doi.org/10.1016/j.dche.2022.100080>
- Pinto, J., Ramos, J. R. C., Costa, R. S., Rossell, S., Dumas, P., & Oliveira, R. (2023). Hybrid deep modeling of a CHO-K1 fed-batch process: combining first-principles with deep neural networks. *Frontiers in Bioengineering and Biotechnology*, 11(September), 1–16. <https://doi.org/10.3389/fbioe.2023.1237963>
- Rashedi, M., Rafiei, M., Demers, M., Khodabandehlou, H., Wang, T., Tulsyan, A., Undey, C., & Garvin, C. (2023a). Machine learning-based model predictive controller design for cell culture processes. *Biotechnology and Bioengineering*, 120(8), 2144–2159. <https://doi.org/10.1002/bit.28486>
- Rashedi, M., Rafiei, M., Demers, M., Khodabandehlou, H., Wang, T., Tulsyan, A., Undey, C., & Garvin, C. (2023b). Machine learning-based model predictive controller design for cell culture processes. *Biotechnology and Bioengineering*, 120(8), 2144–2159. <https://doi.org/10.1002/bit.28486>

- Rashid, M. M., Jackson, R., Parulekar, S. J., & Cinar, A. (2021). Critical quality predictive control of fed-batch mammalian cell bioreactors. *IFAC-PapersOnLine*, 54(3), 330–335. <https://doi.org/10.1016/j.ifacol.2021.08.263>
- Rathore, A. S., Mishra, S., Nikita, S., & Priyanka, P. (2021). Bioprocess control: Current progress and future perspectives. *Life*, 11(6). <https://doi.org/10.3390/life11060557>
- Ratih, R., Asmari, M., Abdel-Megied, A. M., Elbarbry, F., & El Deeb, S. (2021). Biosimilars: Review of regulatory, manufacturing, analytical aspects and beyond. *Microchemical Journal*, 165(March). <https://doi.org/10.1016/j.microc.2021.106143>
- Reddy, J. V., Raudenbush, K., Papoutsakis, E. T., & Ierapetritou, M. (2023). Cell-culture process optimization via model-based predictions of metabolism and protein glycosylation. *Biotechnology Advances*, 67(November 2022), 108179. <https://doi.org/10.1016/j.biotechadv.2023.108179>
- Reyes, S. J., Durocher, Y., Pham, P. L., & Henry, O. (2022). Modern Sensor Tools and Techniques for Monitoring, Controlling, and Improving Cell Culture Processes. *Processes*, 10(2), 189. <https://doi.org/10.3390/pr10020189>
- Rish, A. J., Drennen, J. K., & Anderson, C. A. (2022). Metabolic trends of Chinese hamster ovary cells in biopharmaceutical production under batch and fed-batch conditions. *Biotechnology Progress*, 38(1). <https://doi.org/10.1002/btpr.3220>
- Romanova, N., Niemann, T., Greiner, J. F. W., Kaltschmidt, B., Kaltschmidt, C., & Noll, T. (2021). Hyperosmolality in CHO culture: Effects on cellular behavior and morphology. *Biotechnology and Bioengineering*, 118(6), 2348–2359. <https://doi.org/10.1002/bit.27747>
- Rouiller, Y., Solacroup, T., Deparis, V., Barbaferi, M., Gleixner, R., Broly, H., & Eon-Duval, A. (2012). Application of Quality by Design to the characterization of the cell culture process of an Fc-Fusion protein. *European Journal of Pharmaceutics and Biopharmaceutics*, 81(2), 426–437. <https://doi.org/10.1016/j.ejpb.2012.02.018>
- Rudge, S. R., & Nims, R. W. (2017). ICH Q6B specifications: test procedures and acceptance criteria for biotechnological/biological products. *ICH Quality Guidelines: An Implementation Guide*, 467–486.

- Ruiz, D., Cantón, J., Nougués, J. M., Espuña, A., & Puigjaner, L. (2000). Fault diagnosis system support for reactive scheduling in multipurpose batch chemical plants. *Computer Aided Chemical Engineering*, 8(C), 745–750. [https://doi.org/10.1016/S1570-7946\(00\)80126-1](https://doi.org/10.1016/S1570-7946(00)80126-1)
- Sabatine, M. S., Giugliano, R. P., Keech, A. C., Honarpour, N., Wiviott, S. D., Murphy, S. A., Kuder, J. F., Wang, H., Liu, T., Wasserman, S. M., Sever, P. S., & Pedersen, T. R. (2017). Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. *New England Journal of Medicine*, 376(18), 1713–1722. <https://doi.org/10.1056/nejmoa1615664>
- Sakaki, A., Namatame, T., Nakaya, M., & Omasa, T. (2023). Model-based control system design to manage process parameters in mammalian cell culture for biopharmaceutical manufacturing. *Biotechnology and Bioengineering*, August. <https://doi.org/10.1002/bit.28593>
- Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., & Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling and Software*, 114(January), 29–39. <https://doi.org/10.1016/j.envsoft.2019.01.012>
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2), 259–270. <https://doi.org/10.1016/j.cpc.2009.09.018>
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). Global sensitivity analysis: The primer. In *Global Sensitivity Analysis: The Primer*. <https://doi.org/10.1002/9780470725184>
- Saltelli, A., & Tarantola, S. (2002). On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal. *Journal of the American Statistical Association*, 97(459), 702–709. <https://doi.org/10.1198/016214502388618447>

- Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models* (Vol. 1). Wiley Online Library.
- Sandor, M., Rüdinger, F., Solle, D., Bienert, R., Grimm, C., Groß, S., & Scheper, T. (2013). NIR-spectroscopy for bioprocess monitoring & control. *BMC Proceedings*, 7(S6), 6–8. <https://doi.org/10.1186/1753-6561-7-s6-p29>
- Santos, R. M., Kaiser, P., Menezes, J. C., & Peinado, A. (2019). Improving reliability of Raman spectroscopy for mAb production by upstream processes during bioprocess development stages. *Talanta*, 199(February), 396–406. <https://doi.org/10.1016/j.talanta.2019.02.088>
- Sarna, S., Patel, N., Corbett, B., McCready, C., & Mhaskar, P. (2023). Process-aware data-driven modelling and model predictive control of bioreactor for the production of monoclonal antibodies. *Canadian Journal of Chemical Engineering*, 101(5), 2677–2692. <https://doi.org/10.1002/cjce.24752>
- Schmidt, A., Hengelbrock, A., & Strube, J. (2024). Biopharma 4.0 for Biologics Manufacturing Under Pandemic Constraints. In *Biopharmaceutical Manufacturing: Progress, Trends and Challenges* (pp. 297–317). Springer.
- Schmitt, J., Downey, B., Beller, J., Russell, B., Quach, A., Lyon, D., Curran, M., Mulukutla, B. C., & Chu, C. (2019). Forecasting and control of lactate bifurcation in Chinese hamster ovary cell culture processes. *Biotechnology and Bioengineering*, 116(9), 2223–2235. <https://doi.org/10.1002/bit.27015>
- Schwenzer, M., Ay, M., Bergs, T., & Abel, D. (2021). Review on model predictive control: an engineering perspective. *International Journal of Advanced Manufacturing Technology*, 117(5–6), 1327–1349. <https://doi.org/10.1007/s00170-021-07682-3>
- Sellick, C. A., Croxford, A. S., Maqsood, A. R., Stephens, G., Westerhoff, H. V., Goodacre, R., & Dickson, A. J. (2011). Metabolite profiling of recombinant CHO cells: Designing tailored feeding regimes that enhance recombinant antibody production. *Biotechnology and Bioengineering*, 108(12), 3025–3031. <https://doi.org/10.1002/bit.23269>

- Sendrescu, D. (2011). Nonlinear model predictive control of a depollution bioprocess. *Proceedings - PACCS 2011: 2011 3rd Pacific-Asia Conference on Circuits, Communications and System, 1*, 1–4. <https://doi.org/10.1109/PACCS.2011.5990161>
- Shen, X., & Meinshausen, N. (2023). *Engression: Extrapolation for Nonlinear Regression?* <http://arxiv.org/abs/2307.00835>
- Shirsat, N. P., English, N. J., Glennon, B., & Al-Rubeai, M. (2015). *Modelling of Mammalian Cell Cultures* (Issue October). https://doi.org/10.1007/978-3-319-10320-4_10
- Sidoli, F. R., Mantalaris, A., & Asprey, S. P. (2004). Modelling of mammalian cells and cell culture processes. *Cytotechnology*, 44(1–2), 27–46. <https://doi.org/10.1023/B:CYTO.0000043397.94527.84>
- Silva, F., Resende, D., Amorim, M., & Borges, M. (2020a). *A field study on the impacts of implementing concepts and elements of industry 4.0 in the biopharmaceutical sector*. 1–8. <https://doi.org/10.4995/carma2020.2020.11565>
- Silva, F., Resende, D., Amorim, M., & Borges, M. (2020b). *A field study on the impacts of implementing concepts and elements of industry 4.0 in the biopharmaceutical sector*. 1–8. <https://doi.org/10.4995/carma2020.2020.11565>
- Simon, L. L., Pataki, H., Marosi, G., Meemken, F., Hungerbühler, K., Baiker, A., Tummala, S., Glennon, B., Kuentz, M., Steele, G., Kramer, H. J. M., Rydzak, J. W., Chen, Z., Morris, J., Kjell, F., Singh, R., Gani, R., Gernaey, K. V., Louhi-Kultanen, M., ... Chiu, M. Sen. (2015). Assessment of recent process analytical technology (PAT) trends: A multiauthor review. *Organic Process Research and Development*, 19(1), 3–62. <https://doi.org/10.1021/op500261y>
- Sobol, I. M. (1976). Uniformly distributed sequences with an additional uniform property. *USSR Computational Mathematics and Mathematical Physics*, 16(5), 236–242.
- Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1–3), 271–280.

- Sokolov, M., Soos, M., Neunstoecklin, B., Morbidelli, M., Butté, A., Leardi, R., Solacroup, T., Stettler, M., & Broly, H. (2015). Fingerprint detection and process prediction by multivariate analysis of fed-batch monoclonal antibody cell culture data. *Biotechnology Progress*, 31(6), 1633–1644. <https://doi.org/10.1002/btpr.2174>
- Sommeregger, W., Sissolak, B., Kandra, K., von Stosch, M., Mayer, M., & Striedner, G. (2017). Quality by control: Towards model predictive control of mammalian cell culture bioprocesses. *Biotechnology Journal*, 12(7), 1–7. <https://doi.org/10.1002/biot.201600546>
- Sommerfeld, S., & Strube, J. (2005). Challenges in biotechnology production - Generic processes and process optimization for monoclonal antibodies. *Chemical Engineering and Processing: Process Intensification*, 44(10), 1123–1137. <https://doi.org/10.1016/j.cep.2005.03.006>
- Sou, S. N., Jedrzejewski, P. M., Lee, K., Sellick, C., Polizzi, K. M., & Kontoravdi, C. (2017). Model-based investigation of intracellular processes determining antibody Fc-glycosylation under mild hypothermia. *Biotechnology and Bioengineering*, 114(7), 1570–1582. <https://doi.org/10.1002/bit.26225>
- Sou, S. N., Sellick, C., Lee, K., Mason, A., Kyriakopoulos, S., Polizzi, K. M., & Kontoravdi, C. (2015). How does mild hypothermia affect monoclonal antibody glycosylation? *Biotechnology and Bioengineering*, 112(6), 1165–1176. <https://doi.org/10.1002/bit.25524>
- Spadiut, O., Capone, S., Krainer, F., Glieder, A., & Herwig, C. (2014). Microbials for the production of monoclonal antibodies and antibody fragments. *Trends in Biotechnology*, 32(1), 54–60. <https://doi.org/10.1016/j.tibtech.2013.10.002>
- Steinitz, M. (2014). Human Monoclonal Antibodies. *Methods in Molecular Biology (Clifton, NJ)*, 2014; 1060: 111, 22.
- Strain, B., Morrissey, J., Antonakoudis, A., & Kontoravdi, C. (2023). How reliable are Chinese hamster ovary (CHO) cell genome-scale metabolic models? *Biotechnology and Bioengineering*, 120(9), 2460–2478. <https://doi.org/10.1002/bit.28366>
- Swanson, C. J., Zhang, Y., Dhadda, S., Wang, J., Kaplow, J., Lai, R. Y. K., Lannfelt, L., Bradley, H., Rabe, M., Koyama, A., Reyderman, L., Berry, D. A., Berry, S., Gordon, R.,

- Kramer, L. D., & Cummings, J. L. (2021). A randomized, double-blind, phase 2b proof-of-concept clinical trial in early Alzheimer's disease with lecanemab, an anti-A β protofibril antibody. *Alzheimer's Research and Therapy*, 13(1), 1–14. <https://doi.org/10.1186/s13195-021-00813-8>
- Székely, T., & Burrage, K. (2014). Stochastic simulation in systems biology. *Computational and Structural Biotechnology Journal*, 12(20–21), 14–25. <https://doi.org/10.1016/j.csbj.2014.10.003>
- Tait, A. S., Hogwood, C. E. M., Smales, C. M., & Bracewell, D. G. (2012). Host cell protein dynamics in the supernatant of a mAb producing CHO cell line. *Biotechnology and Bioengineering*, 109(4), 971–982. <https://doi.org/10.1002/bit.24383>
- Tamhane, U. U., & Gurm, H. S. (2008). The chimeric monoclonal antibody abciximab: A systematic review of its safety in contemporary practice. *Expert Opinion on Drug Safety*, 7(6), 809–819. <https://doi.org/10.1517/14740330802500353>
- Tanemura, H., Kitamura, R., Yamada, Y., Hoshino, M., Kakihara, H., & Nonaka, K. (2023). Comprehensive modeling of cell culture profile using Raman spectroscopy and machine learning. *Scientific Reports*, 13(1), 1–13. <https://doi.org/10.1038/s41598-023-49257-0>
- Tang, P., Xu, J., Louey, A., Tan, Z., Yongky, A., Liang, S., Li, Z. J., Weng, Y., & Liu, S. (2020). Kinetic modeling of Chinese hamster ovary cell culture: factors and principles. *Critical Reviews in Biotechnology*, 40(2), 265–281. <https://doi.org/10.1080/07388551.2019.1711015>
- Tangirala, A. K. (2018). *Principles of system identification: theory and practice*. Crc Press.
- Tansey, E. M., & Catterall, P. P. (1995). *Technology transfer in Britain: the case of monoclonal antibodies*.
- Teixeira, A. P., Alves, C., Alves, P. M., Carrondo, M. J. T., & Oliveira, R. (2007). Hybrid elementary flux analysis/nonparametric modeling: Application for bioprocess control. *BMC Bioinformatics*, 8, 1–15. <https://doi.org/10.1186/1471-2105-8-30>

- Teixeira, A. P., Oliveira, R., Alves, P. M., & Carrondo, M. J. T. (2009). Advances in on-line monitoring and control of mammalian cell cultures: Supporting the PAT initiative. *Biotechnology Advances*, 27(6), 726–732. <https://doi.org/10.1016/j.biotechadv.2009.05.003>
- Tenny, M. J., Rawlings, J. B., & Wright, S. J. (2004). Closed-loop behavior of nonlinear model predictive control. *AIChE Journal*, 50(9), 2142–2154. <https://doi.org/10.1002/aic.10177>
- Thompson, M. L., & Kramer, M. A. (1994). Modeling chemical processes using prior knowledge and neural networks. *AIChE Journal*, 40(8), 1328–1340. <https://doi.org/10.1002/aic.690400806>
- Torres, M., Altamirano, C., & Dickson, A. J. (2018). Process and metabolic engineering perspectives of lactate production in mammalian cell cultures. *Current Opinion in Chemical Engineering*, 22, 184–190. <https://doi.org/10.1016/j.coche.2018.10.004>
- United Nations. (2023). *The Sustainable Development Goals: Report 2023*. UN Department of Economic and Social Affairs.
- Vallerio, M., Van Impe, J., & Logist, F. (2014). Tuning of NMPC controllers via multi-objective optimisation. *Computers and Chemical Engineering*, 61, 38–50. <https://doi.org/10.1016/j.compchemeng.2013.10.003>
- Van Beylen, K., Youssef, A., Fernández, A. P., Lambrechts, T., Papantoniou, I., & Aerts, J. M. (2020). Lactate-based model predictive control strategy of cell growth for cell therapy applications. *Bioengineering*, 7(3), 1–18. <https://doi.org/10.3390/bioengineering7030078>
- van den Bosch, P. P. J., & van der Klauw, A. C. (2020). *Modeling, identification and simulation of dynamical systems*. crc Press.
- Vann, L., & Sheppard, J. (2017). Use of near-infrared spectroscopy (NIRs) in the biopharmaceutical industry for real-time determination of critical process parameters and integration of advanced feedback control strategies using MIDUS control. *Journal of Industrial Microbiology and Biotechnology*, 44(12), 1589–1603. <https://doi.org/10.1007/s10295-017-1984-2>

- Vargason, A. M., Anselmo, A. C., & Mitragotri, S. (2021). The evolution of commercial drug delivery technologies. *Nature Biomedical Engineering*, 5(9), 951–967. <https://doi.org/10.1038/s41551-021-00698-w>
- Villiger, T. K., Scibona, E., Stettler, M., Broly, H., Morbidelli, M., & Soos, M. (2016). Controlling the time evolution of mAb N-linked glycosylation - Part II: Model-based predictions. *Biotechnology Progress*, 32(5), 1135–1148. <https://doi.org/10.1002/btpr.2315>
- Villiger, T. K., Steinhoff, R. F., Ivarsson, M., Solacroup, T., Stettler, M., Broly, H., Krismer, J., Pabst, M., Zenobi, R., Morbidelli, M., & Soos, M. (2016). High-throughput profiling of nucleotides and nucleotide sugars to evaluate their impact on antibody N-glycosylation. *Journal of Biotechnology*, 229, 3–12. <https://doi.org/10.1016/j.jbiotec.2016.04.039>
- von Stosch, M., Davy, S., Francois, K., Galvanauskas, V., Hamelink, J. M., Luebbert, A., Mayer, M., Oliveira, R., O’Kennedy, R., Rice, P., & Glassey, J. (2014). Hybrid modeling for quality by design and PAT-benefits and challenges of applications in biopharmaceutical industry. *Biotechnology Journal*, 9(6), 719–726. <https://doi.org/10.1002/biot.201300385>
- von Stosch, M., Portela, R. M., & Varsakelis, C. (2021). A roadmap to AI-driven in silico process development: bioprocessing 4.0 in practice. *Current Opinion in Chemical Engineering*, 33, 100692. <https://doi.org/10.1016/j.coche.2021.100692>
- Walsh, G., & Walsh, E. (2022). Biopharmaceutical benchmarks 2022. *Nature Biotechnology*, 40(12), 1722–1760. <https://doi.org/10.1038/s41587-022-01582-x>
- Walsh, I., Myint, M., Nguyen-Khuong, T., Ho, Y. S., Ng, S. K., & Lakshmanan, M. (2022). Harnessing the potential of machine learning for advancing “Quality by Design” in biomanufacturing. *MAbs*, 14(1). <https://doi.org/10.1080/19420862.2021.2013593>
- Walther, J., Lu, J., Hollenbach, M., Yu, M., Hwang, C., McLarty, J., & Brower, K. (2019). Perfusion Cell Culture Decreases Process and Product Heterogeneity in a Head-to-Head Comparison With Fed-Batch. *Biotechnology Journal*, 14(2), 1–10. <https://doi.org/10.1002/biot.201700733>

- Webster, T. A., Hadley, B. C., Dickson, M., Busa, J. K., Jaques, C., & Mason, C. (2021). Feedback control of two supplemental feeds during fed-batch culture on a platform process using inline Raman models for glucose and phenylalanine concentration. *Bioprocess and Biosystems Engineering*, 44(1), 127–140. <https://doi.org/10.1007/s00449-020-02429-y>
- Webster, T. A., Hadley, B. C., Dickson, M., Hodgkins, J., Olin, M., Wolnick, N., Armstrong, J., Mason, C., & Downey, B. (2023). Automated Raman feed-back control of multiple supplemental feeds to enable an intensified high inoculation density fed-batch platform process. *Bioprocess and Biosystems Engineering*, 46(10), 1457–1470. <https://doi.org/10.1007/s00449-023-02912-2>
- Weishuhn, M. (2024a). *Inciteful: Citation network exploration*. <https://inciteful.xyz/p?ids%5B%5D=W2918163347&ids%5B%5D=W3024508204&ids%5B%5D=W3126869566&ids%5B%5D=W3082333662&ids%5B%5D=W4283017114&ids%5B%5D=W4220673939&ids%5B%5D=W4309159259&ids%5B%5D=W4310760636&ids%5B%5D=W2955350610>
- Weishuhn, M. (2024b). *Inciteful: Citation network exploration*. <https://inciteful.xyz/p?ids%5B%5D=W1870974274&ids%5B%5D=W2021199190&ids%5B%5D=W2040813538&ids%5B%5D=W2606676378&ids%5B%5D=W2911651053&ids%5B%5D=W2918163347&ids%5B%5D=W2944297854&ids%5B%5D=W2998669422&ids%5B%5D=W3081409172&ids%5B%5D=W3089053125&ids%5B%5D=W>
- William Y. Svrcek, Donald P. Mahoney, B. R. Y. (2014). *A Real-Time Approach to Process Control, Third Edition*. 362.
- World Bank, & OECD. (2022). *Gross Domestic Product by Country*. <https://data.worldbank.org/indicator/NY.GDP.MKTP.CD>
- World Health Organization. (2020). *The top 10 causes of death*. <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>

- Xing, Z., Bishop, N., Leister, K., & Li, Z. J. (2010). Modeling kinetics of a large-scale fed-batch CHO cell culture by markov chain monte carlo method. *Biotechnology Progress*, 26(1), 208–219. <https://doi.org/10.1002/btpr.284>
- Xing, Z., Li, Z., Chow, V., & Lee, S. S. (2008). Identifying inhibitory threshold values of repressing metabolites in CHO cell culture using multivariate analysis methods. *Biotechnology Progress*, 24(3), 675–683. <https://doi.org/10.1021/bp070466m>
- Xu, J., Ou, J., McHugh, K. P., Borys, M. C., & Khetan, A. (2022). Upstream cell culture process characterization and in-process control strategy development at pandemic speed. *MAbs*, 14(1). <https://doi.org/10.1080/19420862.2022.2060724>
- Xu, W. J., Lin, Y., Mi, C. L., Pang, J. Y., & Wang, T. Y. (2023). Progress in fed-batch culture for recombinant protein production in CHO cells. *Applied Microbiology and Biotechnology*, 107(4), 1063–1075. <https://doi.org/10.1007/s00253-022-12342-x>
- Yan, X., Dong, X., Wan, Y., Gao, D., Chen, Z., Zhang, Y., Zheng, Z., Chen, K., Jiao, J., Sun, Y., He, Z., Nie, L., Fan, X., Wang, H., & Qu, H. (2024). Development of an in-line Raman analytical method for commercial-scale CHO cell culture process monitoring: Influence of measurement channels and batch number on model performance. *Biotechnology Journal*, December 2023, 1–10. <https://doi.org/10.1002/biot.202300395>
- Yang, O., & Ierapetritou, M. (2021). Mab production modeling and design space evaluation including glycosylation process. *Processes*, 9(2), 1–19. <https://doi.org/10.3390/pr9020324>
- Yang, W. C., Minkler, D. F., Kshirsagar, R., Ryll, T., & Huang, Y. M. (2016). Concentrated fed-batch cell culture increases manufacturing capacity without additional volumetric capacity. *Journal of Biotechnology*, 217, 1–11. <https://doi.org/10.1016/j.jbiotec.2015.10.009>
- Yatipanthala, B. S., & Gras, S. L. (2023). Predictive models for upstream mammalian cell culture development - A review. *Digital Chemical Engineering*, 10(December 2023), 100137. <https://doi.org/10.1016/j.dche.2023.100137>

- Yee, J. C., Rehmann, M. S., Yao, G., Sowa, S. W., Aron, K. L., Tian, J., Borys, M. C., & Li, Z. J. (2018). Advances in process control strategies for mammalian fed-batch cultures. *Current Opinion in Chemical Engineering*, 22(Figure 1), 34–41. <https://doi.org/10.1016/j.coche.2018.09.002>
- Zahavi, D., & Weiner, L. (2020). Monoclonal Antibodies in Cancer Therapy. *Antibodies*, 9(3), 34. <https://doi.org/10.3390/antib9030034>
- Zakrzewski, R., Lee, K., & Lye, G. J. (2022). Development of a miniature bioreactor model to study the impact of pH and DOT fluctuations on CHO cell culture performance as a tool to understanding heterogeneity effects at large-scale. *Biotechnology Progress*, 38(4), 1–15. <https://doi.org/10.1002/btpr.3264>
- Zavala-Ortiz, D. A., Ebel, B., Li, M. Y., Barradas-Dermitz, D. M., Hayward-Jones, P. M., Aguilar-Uscanga, M. G., Marc, A., & Guedon, E. (2020). Support Vector and Locally Weighted regressions to monitor monoclonal antibody glycosylation during CHO cell culture processes, an enhanced alternative to Partial Least Squares regression. *Biochemical Engineering Journal*, 154(November 2019), 107457. <https://doi.org/10.1016/j.bej.2019.107457>
- Zhang, L., & Mao, S. (2017). Application of quality by design in the current drug development. *Asian Journal of Pharmaceutical Sciences*, 12(1), 1–8. <https://doi.org/10.1016/j.ajps.2016.07.006>
- Zhang, X. Y., Trame, M. N., Lesko, L. J., & Schmidt, S. (2015). Sobol sensitivity analysis: A tool to guide the development and evaluation of systems pharmacology models. *CPT: Pharmacometrics and Systems Pharmacology*, 4(2), 69–79. <https://doi.org/10.1002/psp4.6>
- Zhu, M. M., Goyal, A., Rank, D. L., Gupta, S. K., Vanden Boom, T., & Lee, S. S. (2005). Effects of elevated pCO₂ and osmolality on growth of CHO cells and production of antibody-fusion protein B1: A case study. *Biotechnology Progress*, 21(1), 70–77. <https://doi.org/10.1021/bp049815s>

- Zhu, Z., Chen, X., Li, W., Zhuang, Y., Zhao, Y., & Wang, G. (2023). Understanding the effect of temperature downshift on CHO cell growth, antibody titer and product quality by intracellular metabolite profiling and in vivo monitoring of redox state. *Biotechnology Progress*, 39(4), 1–20. <https://doi.org/10.1002/btpr.3352>
- Zobel-Roos, S., Schmidt, A., Mestmäcker, F., Mouellef, M., Huter, M., Uhlenbrock, L., Kornecki, M., Lohmann, L., Ditz, R., & Strube, J. (2019). Accelerating biologics manufacturing by modeling or: Is Approval under the QbD and PAT approaches demanded by authorities acceptable without a digital-twin? *Processes*, 7(2), 1–28. <https://doi.org/10.3390/pr7020094>
- Zupke, C., Brady, L. J., Slade, P. G., Clark, P., Caspary, R. G., Livingston, B., Taylor, L., Bigham, K., Morris, A. E., & Bailey, R. W. (2015). Real-time product attribute control to manufacture antibodies with defined N-linked glycan levels. *Biotechnology Progress*, 31(5), 1433–1441. <https://doi.org/10.1002/btpr.2136>