Uncertainty-aware decision-support tools for next-generation pharmaceutical supply chains

Miriam Sarkis

November 2024

A thesis submitted for the degree of Doctor of Philosophy and the Diploma of Membership of Imperial College London

The Sargent Centre for Process Systems Engineering
Department of Chemical Engineering
Imperial College London, United Kingdom

Declaration of originality

I hereby certify that this is my own work and any material that is not my own work has been appropriately acknowledged and referenced. No portion of the work presented has been submitted in support of another degree or qualification at any other university or institute of learning.

Copyright statement

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.

Abstract

Supply chain resilience and sustainability are becoming growing priorities in the pharmaceutical sector. In recent years, the industry has seen a market boom of next generation therapies allowing treatment of otherwise incurable diseases and platform vaccine technologies. Manufacturers catering for these markets reported shortages and delays due to unforeseen demand growth combined with uncertainty in manufacturing capabilities of platforms still under development. At the same time, regulatory and industrial initiatives are urging manufacturers to quantify and reduce supply chain environmental footprints. This highlights an opportunity for computer-aided modelling and optimisation to support investment planning and help quantify trade-offs between cost, environmental, social and resilience impact indicators.

This thesis presents decision-support tools that assess resilience and sustainability *a priori* for supply chain networks require rapid scale-up and operate under supply-demand uncertainty. Firstly, the cost optimisation problem under process uncertainty is considered. In this context, techno-economic modelling and uncertainty analysis are used to quantify manufacturing uncertainties necessary to mitigate against when proceeding in investment planning through stochastic optimisation. Leveraging on the acquired knowledge on process uncertainty, a feasibility problem is formulated which helps quantify an uncertainty-aware flexibility space for a given design. Secondly, the cost optimisation problem under demand and process uncertainty is considered to identify planning strategies that flexibly adapt to time-dependent uncertainties. Finally, the integration of sustainability-related metrics in the investment planning problem is proposed. It is demonstrated that given a target demand, cost minimisation reduces environmental footprint for the case study considered. To conclude, key findings are discussed with an outlook on emerging research directions to combine capabilities in integrated decision-support platforms.

Acknowledgements

I would like to thank my supervisors Dr. Maria Papathanasiou, for believing in me and going above and beyond as a mentor, supporting me in critical times and celebrating every success, and Prof. Nilay Shah, for the support, guidance and volume of knowledge shared which inspires me to this day.

I would like to thank Prof. Cleo Kontoravdi, Prof. Suzanne Farid and Dr. Andrea Bernardi for the valuable feedback and knowledge they shared, helping me develop my PhD project, Prof. Claire Adjiman and Prof. Ronny Pini for the insightful advice given before the start, during and towards the end of it, and Prof. Ruth Misener and Prof. Ana Barbosa-Póvoa for the time dedicated to examine this thesis and the expertise shared.

I would like to thank Dr. Zoltan Kis and Dr. Kyungjae Tak for their insights in techno-economic and sensitivity analysis and the students Jesslyn, Hale, Isabel, Kye, Alasdair and Chung-Lin, for their hard work which opened avenues within my project in the area of sustainability.

I would like to thank my lab group, Steven, Niki, Foteini, Daniele, Alexandros and Andrea, for creating a special space to share ideas and desserts and always having each other's back throughout the research journey and life.

I would like to thank Mariana, Nikki, Tanuj, Tom, Simon, Ilya, Gustavo, Shubhani, Niki, Ahmed, Céline, Lucy, Mathias, Thomas, Ben, James, Kostis, Gaby, Lilly, Zeynep, Miguel, Tom, Emma, Francesco, Ben, Max, Deniz, Andrea and Ethan for the lunch and coffee breaks, the scientific conversations inspiring me at unexpected times, the conferences together and all the fun, especially on Thursday and Friday evening, and Jana and Ieke for the chocolate chats.

I would like to thank Aditi, Ben, Carmen, Ravi, Ioanna, Alexis, Komal, Hassan, Natsuko, Camilla and Marvi for listening to my supply chain monologues and for their timeless friendship, and the ICVC family, for the Tuesday evenings, Wednesday afternoons and Sundays spent to rest the brain and grow as a team.

Last but not least, I would like to thank my parents and my brothers, Gabriele and Davide, for their unwavering support and love. I would like to thank my mum, Elena, for being near me daily on the phone on my way to and from work and my father, Claude, for thinking ahead and helping me shape my future, always encouraging me to go my way and do my best.

Contents

List of Symbols **vii**

List of Figures **xiii**

List of Tables **xviii**

1 Introduction **1**

1.1 Pharmaceutical industrial landscape 2

1.2 Aims & objectives 6

2 Literature review **8**

2.1 Introduction 9

2.2 Performance quantification 10

2.2.1 Techno-economic modelling 10

2.2.2 Life cycle assessment 12

2.2.3 Remarks 13

2.3 Multi-scale modelling 13

2.3.1 Scheduling 13

2.3.2 Capacity planning 17

2.3.3 Remarks 22

2.4 Planning under uncertainty 23

2.4.1 Optimisation under uncertainty 23

2.4.2 Simulation-optimisation 28

2.4.3 Remarks 32

2.5 Multi-objective optimisation 33

2.5.1 Emerging trade-offs 33

2.5.2 Remarks 35

2.6 Contribution of this work 35

3 Quantification of process uncertainty **37**
3.1 Introduction . . . 38
3.2 Methodology . . . 39
3.2.1 Data collection . . . 40
3.2.2 Flowsheet development . . . 40
3.2.3 Uncertainty analysis . . . 50
3.2.4 Global sensitivity analysis . . . 51
3.3 Results & discussion . . . 53
3.3.1 Techno-economic analysis of vector processes . . . 53
3.3.2 Uncertainty quantification . . . 55
3.3.3 Uncertainty characterisation . . . 60
3.3.4 Supply chain implications . . . 65
3.4 Conclusions . . . 67

4 Integration of process uncertainty **68**
4.1 Introduction . . . 69
4.2 Methodology . . . 71
4.2.1 Problem statement . . . 71
4.2.2 Process scenario generation . . . 73
4.2.3 Optimisation formulation . . . 75
4.2.4 Stochastic simulation framework . . . 83
4.2.5 Computational resources . . . 84
4.2.6 Case study description . . . 86
4.3 Results & discussion . . . 86
4.3.1 Optimisation results . . . 86
4.3.2 Simulation results . . . 91
4.3.3 Scenario tuning . . . 94
4.4 Conclusions . . . 96

5 Quantification of flexibility under uncertainty **98**
5.1 Introduction . . . 99
5.2 Methodology . . . 101
5.2.1 Problem statement . . . 101
5.2.2 Process scenario generation . . . 103
5.2.3 Optimisation formulation . . . 104

5.2.4 Stochastic simulation framework . . . 113
5.2.5 Case study description . . . 114
5.2.6 Computational resources . . . 114
5.3 Results & discussion . . . 116
5.3.1 Optimisation . . . 116
5.3.2 Simulation . . . 121
5.4 Conclusions . . . 125

6 Integration of demand uncertainty 127
6.1 Introduction . . . 128
6.2 Methodology . . . 130
6.2.1 Problem statement . . . 130
6.2.2 Techno-economic performance quantification . . . 131
6.2.3 Optimisation formulation . . . 132
6.2.4 Case studies . . . 143
6.2.5 Model statistics . . . 145
6.3 Results & discussion . . . 146
6.3.1 Techno-economic analysis . . . 146
6.3.2 Investment planning under demand uncertainty . . . 148
6.3.3 Integration of process uncertainty . . . 151
6.4 Conclusion . . . 155

7 Integration of sustainability metrics 157
7.1 Introduction . . . 158
7.2 Methodology . . . 160
7.2.1 Problem statement . . . 160
7.2.2 Environmental footprint quantification . . . 162
7.2.3 Optimisation formulation . . . 163
7.2.4 Computational resources . . . 171
7.2.5 Uncertainty quantification . . . 172
7.3 Results & discussion . . . 172
7.3.1 Quantification of environmental footprint . . . 172
7.3.2 Deterministic optimisation results . . . 176
7.3.3 Quantification of footprints uncertainty . . . 180
7.4 Conclusions . . . 182

8 Conclusions & outlook **184**
8.1 Summary of contributions . . . 185
8.2 Main conclusions . . . 187
8.3 Future directions . . . 188
8.3.1 Investment planning . . . 189
8.3.2 Flexibility metrics quantification . . . 189
8.3.3 *Uncertainty-aware* and sustainable designs . . . 190
8.3.4 Other applications . . . 191

Dissemination of outputs **192**

A Techno-economic modelling **216**
A.1 Materials . . . 217
A.1.1 AdV manufacturing . . . 217
A.1.2 LV manufacturing . . . 220
A.1.3 AAV manufacturing . . . 223
A.2 Mixed equipment flowsheets . . . 224
A.2.1 AdV manufacturing . . . 224
A.2.2 LV manufacturing . . . 226
A.2.3 AAV manufacturing . . . 228
A.3 MU equipment flowsheets . . . 231
A.3.1 AdV manufacturing . . . 231
A.3.2 LV manufacturing . . . 233
A.4 SU equipment flowsheets . . . 235
A.4.1 LV manufacturing . . . 235
A.4.2 AdV manufacturing . . . 235

B Global sensitivity analysis **240**
B.1 GSA I & GSA II results . . . 241

C Optimisation inputs **248**
C.1 Global supply chain parameters . . . 249
C.2 Parameters Chapter 4 and 5 . . . 250
C.2.1 Deterministic cases . . . 250
C.2.2 Stochastic cases . . . 252
C.3 Parameters Chapter 6 . . . 254

C.3.1 MU case 255
C.3.2 SU case 257
C.4 Parameters Chapter 7 257
C.4.1 Costs 257
C.4.2 Impacts 260

D Life cycle assessment modelling **265**
D.1 Stainless steel equipment 266
D.2 Plastics-based equipment 267
D.3 LCA inputs and outputs 268

List of Symbols

Acronyms

AAV	Adeno-associated viral vector
ADP	Approximate dynamic programming
AdV	Adenoviral vector
ATMP	Advanced Therapy Medicinal Product
C&GT	Cell & gene therapies
CAR	Chimeric antigen receptor (T-cells)
CIP	Cleaning-in-place
COGs	Cost of goods
DALY	Disability-adjusted-life-year
DSP	Downstream process
Eco	Ecosystems, species/yr
F&F	Fill & finish process
FEC	Freshwater eco-toxicity, kg1,4-DCB
FEU	Freshwater eutrophication, kg Peq
FPMF	Fine particle material formation, kg PM2.5 eq
FRS	Fossil resource scarcity, kg oil eq
GSA	Global sensitivity analysis
GWP	Global warming potential, kg CO_2 eq
HC	Human carcinogenic toxicity, kg 1,4-DCB
HH	Human health, DALY
HNC	Human non-carcinogenic toxicity, kg 1,4-DCB

IEX Ion exchange chromatography

IQR Inter-quartile range

IR Ionising radiation, kBq Co-60 eq

KPI Key Performance Indicator

LCA Life Cycle Assessment

LCI Life Cycle Inventory

LSA Local sensitivity analysis

LU Land use, m^2a crop eq

LV Lentiviral vector

mAb Monoclonal antibody

MC Monte Carlo

MECO Marine ecotoxicity, kg 1,4-DCB

MEU Marine eutrophication, kg N eq

MILP Mixed integer linear program

mp Multiparametric

MPC Model predictive control

MRS Mineral resource scarcity, kg Cu eq

MU Multi-use

NPV Net present value

OFHH Ozone formation, human health, kg NOx eq

OFTE Ozone formation, terrestrial ecosystems, kg NOx eq

PCA Principal component analysis

PMI Process mass intensities

PSE Process systems engineering

QALY Quality-adjusted-life-year

QC/QA Quality control/quality assurance

QMC Quasi-Monte Carlo

R&D Reseach and development

RA Resource availability, USD2013

RS-HDMR Random sampling high dimensional model representation

RTN Resource-task network

SAA Sample average approximation

SIP Steam-in-place

SOD Stratospheric ozone depletion, kg CFC-11 eq

STN State-task network

SU Single-use

SUT Single-use technologies

TAP Terrestrial acidification, kg SO_2 eq

TETP Terrestrial ecotoxicity, kg 1,4-DCB

UA Uncertainty analysis

UF/DF Ultrafiltration/diafiltration

USP Upstream process

VV Viral vector

WC Water consumption, m^3

WFI Water for injection

Binary variables

H 1 if production line is installed, 0 otherwise

M 1 if product-specific equipment is installed, 0 otherwise

V 1 if scale is selected, 0 otherwise

W 1 if product allocation occurs, 0 otherwise

X 1 if supply chain link is established, 0 otherwise

Y 1 if node is established, 0 otherwise

Z 1 if production line is available, 0 otherwise

Continuous/Integer Variables

δ converged output level, %

σ output level δ in MC run, %

B number of batches processed, batches

C^{EQj} cost of equipment investment in plant, \$/y

C^{FACj} annual facility-dfendent cost, \$/y

C^{INSj} cost of plant installation, \$/y

C^{LABj} annual labour cost, \$/y

C^{VARj} annual variable cost, \$/y

E^{EP} total endpoint impact, midpoint units

E^{MP} total midpoint impact, midpoint units

G shortage, batches

I supply chain inventory, batches

J manufacturing inventory, batches

P production level, batches

Q number of batches shipped, batches

Q^{IN} inbound flow to node, batches

Q^{OUT} outbound flow from node, batches

S consumption (i.e. sales), batches

T time allocated to production, days

TC^{CAPj} total capital cost of manufacturing, \$

TC^{OPj} total operating cost of manufacturing, \$/y

TC^{OPk} total operating cost of storage, \$/y

TC^{T} total cost of transport, \$/y

TE total monetised environmental footprint, USD2017

U utilisation of production line, days

z total supply chain cost, \$

Indices

a scale

e equipment type

f endpoint

i	product
l	manufacturing line
n, m	node
p, k	process scenario
t	time period
w	midpoint
y	demand scenario

Other

c^{CAP}	capital cost output (i.e. c^{EQj}), \$
c^{OP}	operating cost output (i.e. c^{LABj}, c^{VARj}, c^{FACj}), \$
C_1	direct fixed capital cost, \$
C_2	start-up and validation cost, \$
C_3	working capital cost, \$
O_1	materials cost, \$
O_2	labour cost, \$
O_3	facility-dependent cost, \$
O_4	consumables cost, \$
O_5	utilities cost, \$
S_T	total order sensitivity index, -
S_{ij}	second order sensitivity index, -
S_i	first order sensitivity index, -

Parameters

α	process time, days
χ	batch size doses/batch
Γ_ξ	matrix of values for each general parameter ξ of dimensions $\Omega \times d$
μ	monetisation factor of endpoint, USD2017/endpoint units
Ω	number of samples generated via UA, -
ω	number of samples to reduce Ω, -

ϕ process scenario probability, -

ψ shortage penalty, $

ρ transport distance, km

τ bottleneck time, days

Θ matrix of input process parameters of dimensions $\Omega \times d$

θ submatrix of Θ of dimensions $\omega \times d$

ξ general process parameters, -

c^{EQjl} cost of manufacturing line equipment investment, $/line

c^{FACjl} facility cost per line, $/line

c^{LABjl} labour cost per batch, $/batch

c^{OPk} cost of storage per batch, $/batch

c^{T} cost of transport parameter, $/batch/km

c^{VARjl} variable cost per batch, $/batch

D demand in demand zone, doses

D^{T} total demand, doses

e^{EP} endpoint impact per batch, endpoint units

e^{MP} midpoint impact per batch, midpoint units

K number of simulation scenarios

O number of optimisation scenarios

Q^{MAX} upper bound batch flow, batches

Q^{MIN} lower bound batch flow, batches

R number of Monte Carlo runs

T^{MAX} upper bound time allocation, days

T^{MIN} lower bound time allocation, days

U^{MAX} upper bound utilisation, days

X_{random} matrix of random integers

X_{Sobol} matrix of quasi-random integers

X_{test} matrix of random integers

Sets

A	set of scales: $\{50, 200, 1000, 2000\}$
d	set of sampling dimensions: e.g. $A \times I$
EP	set of endpoints: $\{HH, RA, Eco\}$
EQ	set of equipment types: $\{MU, SU\}$
I	set of products: $\{LV, AdV\}$
L^S	set of lines within section s: $\{L_1, L_2, L_3, L_4\}$
M^N	set of downstream nodes w.r.t. node n
MP	set of midpoints: $\{GW, FEU, WC...\}$
N^J	set of manufacturing nodes: $\{J_1, J_2, J_3\}$
N^K	set of storage nodes: $\{K_1\}$
N^M	set of upstream nodes w.r.t. node m
N^Z	set of demand zones: $\{Z_1, Z_2, Z_3, Z_4, Z_5\}$
P	set of demand scenarios: $\{DS_1, DS_2, ..., DS_N\}$
P	set of process scenarios: $\{P_1, P_2, ..., P_N\}$ (stochastic) or $\{P_1\}$ (deterministic)
P^O	set of process scenarios for optimisation: $\{P_1, P_2, ..., P_O\}$
P^S	set of process scenarios for simulation: $\{P_1, P_2, ..., P_K\}$
S	set of process sections: $\{U, D, F\}$
T	set of time periods: $\{T_1, T_2, ..., T_N\}$

List of Figures

1.1 General schematic of viral vector supply chains, comprising (1) primary manufacturing, (2) secondary manufacturing/fill-to-finish, (3) centralised storage, (4) decentralised storage, (5) hospitals or administration points as well as (7) specialised leukapheresis sites, (6) cell therapy manufacturing facilities. In this work, we focus on the complexities of viral vector manufacturing for a range of applications and supply chain planning for vaccines and *ex-vivo* gene therapy (*System*). 5

2.1 System representation in capacity planning problems. General approaches: (a) discrete set of available sizes, (b) capacity increments (Δ). Approaches based on pharmaceutical industry facilities: (c) parallel suites comprising a production and purification line (Rotstein et al., 1999; Papageorgiou et al., 2001), (d) decoupled upstream and downstream suites (Gatica et al., 2003a). . . . 19

2.2 Network design: active nodes within P_N, F_N, S_N, D_N allow links and shipments 20

2.3 Problem representations for optimisation under uncertainty, specifically stochastic programming and robust optimisation: (a) Two-stage, (b) two-stage multi-period, (c) multi-stage multi-period with exogenous uncertainty (with realisations at $t = 2, 4, 6$), (d) multi-stage multi-period with endogenous uncertainty (with realisations timings depending on the decisions made at each stage). . 24

3.1 Flowsheet for: (a) Adv manufacturing highlighting labeled process steps (1–15) and key resource inputs, (b) LV manufacturing highlighting labeled process steps (1–13) and key resource inputs, and (c) AAV manufacturing highlighting labeled process steps (1–15) and key resource inputs. 41

3.2 (a) Overview of layers of calculations embedded in techno-economic model and input-output space overview for GSA and UA and (b) Sequential GSA workflow. Step 1: highest-ranking ($S_i + S_{ij} \geq 0.05$) operational inputs are obtained, which impact throughput-related outputs. Step 2: the screen operational inputs are combined with resource costs inputs. GSA determines the highest-ranking inputs within the operational and resource costs input space. Outputs monitored in GSA II include capital costs and operating costs. Finally, GSA III is performed. 50

3.3 Breakdown of costs of manufacturing for viral vector processes: (a) capital costs for AdV process, (b) operating costs for AdV process, (c) capital costs for LV process, (d) operating costs for LV process, (e) capital costs for AAV process, (f) operating costs for AAV process. 54

3.4 Violins plots for throughput related quantities comparing two input distributions including uniform and triangular. Scale-dependent results are shown for (a) process time for AdV, (b) process time for LV, (c) process time AAV, (e) cycle time for AdV, (e) cycle time for LV, (f) cycle time for AAV, (g) batch size for AdV, (h) batch size for LV, (i) batch size for AAV. 56

3.5 Violins plots for cost related quantities comparing two input distributions including uniform and triangular. Scale-dependent results are shown for (a) Capital cost for adenoviral (AdV), (b) capital cost for lentiviral (LV), (c) capital cost for adeno-associated vector (AAV), (d) operating cost for AdV, (e) operating cost for LV, (f) operating cost for AAV, (g) unit cost for AdV, (h) unit cost for LV, (i) unit cost for AAV. 59

3.6 Global sensitivity analysis (GSA) plots for throughput-related outputs of each viral vector platforms and highest ranking inputs. Each plot displays first-order sensitivity indices S_i along the diagonal and second-order indices S_{ij} in squares combining two inputs: (a) process time AdV, (b) process time for LV, (c) process time for AAV, (e) cycle time for AdV, (e) cycle time for LV, (f) cycle time for AAV, (g) batch size for AdV, (h) batch size for LV, (i) batch size for AAV. 61

3.7 Global sensitivity analysis (GSA) plots for cost-related outputs of each viral vector platforms displaying highest ranking inputs. Each plot displays first-order sensitivity indices S_i along the diagonal and second-order indices S_{ij} in squares combining two inputs: (a) capital cost for AdV, (b) capital cost for LV, (c) capital cost for AAV, (d) operating cost for AdV, (e) operating cost for LV, (f) operating cost for AAV, (g) unit cost for AdV, (h) unit cost for LV, (i) unit cost for AAV. 64

4.1 Optimisation-simulation framework: (a) process scenarios are generated for the optimisation (b) the optimisation is formulated as a deterministic problem or a stochastic program, (c) the simulation scheme is implemented to quantify the statistical performance of each optimisation solution. 71

4.2 Superstructure for optimisation of single-product (LV) case study. Nodes n comprise manufacturing nodes (square), storage (triangle) and demand zones (circles). Each manufacturing node is modelled as a set of process sections s comprising parallel manufacturing lines l at scale a. Process scenarios p are generated as input datasets for the optimisation-simulation framework. 73

4.3 Network KPIs for single-product case: 3 different demand realisations (low, medium, high) and 4 different process realisations (best-case, median case, worst-case, stochastic-case) for LV (a-f) . 88

4.4 Supply chain structures and product allocation for different optimisation strategies and demands. *Best* case (a) low, (b) mid, (c) high demand, *median* case (d) low, (e) mid, (f) high demand, *worst* case (g) low, (h) mid, (i) high demand, *quasirandom* case (j) low, (k) mid, (l) high demand, *random* case (m) low, (n) mid, (o) high demand. Product allocated (■), inactive lines (□). 89
4.5 Simulation via random sampling ($P^S = 2^6$ samples) for single-product network structures generated for (a) low, (b) mid, (c) high demand for LV and cost per dose for (d) low, (e) mid, (f) high demand. 92
4.6 Multiple-runs simulation for different designs obtained with optimisation strategies: (a) mean probability of meeting target demand, (b) mean probability of cost below target 94
4.7 Scenario tuning results for Trial 1 and demand (a) low, (c) mid, (e) high and Trial 2 and demand (b) low, (d) mid, (f) high. 95

5.1 Optimisation-simulation framework: (a) process scenarios are generated for the optimisation (b) the optimisation is formulated as a deterministic problem or a stochastic program, (c) the simulation scheme is implemented to quantify the statistical performance of each optimisation solution. 101
5.2 Superstructure for optimisation of multi-product i case study. Nodes n comprise manufacturing nodes (square), storage (triangle) and demand zones (circles). Each manufacturing node is modelled as a set of process sections s comprising parallel manufacturing lines l at scale a. Process scenarios p are generated as input datasets for the optimisation-simulation framework. 103
5.3 Networks for unconstrained case (a-f): AdV production (■), LV production (■), multi-product production, (■), uninstalled inactive lines (□). 117
5.4 Cost performance unconstrained case: (a) capital cost, (b) operating cost 117
5.5 Network structures for constraints case (g-l): AdV production (■), LV production (■), multi-product production, (■), installed inactive lines (■), uninstalled inactive lines (□). 120

5.6 Cost performance constrained case: (a) capital cost, (b) operating cost . . . 120
5.7 Simulations network (a-f), for unconstrained case . . . 123
5.8 Multiple runs simulations for network (a-f), for unconstrained case 123
5.9 Simulations network g to l, for constrained case . . . 124
5.10 Multiple runs simulations for network (g-l), for constrained case . 124

6.1 Superstructure for optimisation of multi-product i case study. Nodes n comprise manufacturing nodes (square), storage (triangle) and demand zones (circles). Each manufacturing node is modelled as a set of process sections s comprising parallel manufacturing lines l at scale a, according to demand scenarios y at time t. . . . 132
6.2 Demand scenarios: (a) LV demand tree, (b) AdV demand tree. The optimisation considers a joint scenario tree based on these outcomes for each product. . . . 144
6.3 Cost breakdown for SU vs MU manufacturing platforms for LV manufacturing at (a) 50L, (b) 200L, (c) 1000L, (d) 2000L scale and AdV manufacturing at (e) 50L, (f) 200L, (g) 1000L and (h) 2000L scale. . . . 147
6.4 Multistage optimisation structures for deterministic case for MU case (a-d) and SU case (e-h). AdV product allocation (■), LV product allocation (■), inactive installed lines (■), inactive uninstalled lines (□) . . . 149
6.5 Multistage optimisation KPIs for deterministic case for MU case (a-e) and SU case (f-j) . . . 150
6.6 Multistage optimisation structures for initial worst-case for MU case (a-d) and SU case (e-h). AdV product allocation (■), LV product allocation (■), inactive installed lines (■), inactive uninstalled lines (□) . . . 152
6.7 Multistage optimisation initial worst-case results for MU case (a-e) and SU case (f-j) . . . 153

7.1 Quantification-optimisation framework: (a) techno-economic quantification (b) environmental footprint quantification, (c) performance optimisation. . . . 160

7.2 Superstructure for optimisation of LV-based case study. Nodes n comprise manufacturing nodes (square), storage (triangle) and demand zones (circles). Each manufacturing node is modelled as a set of process sections s comprising parallel manufacturing lines l at scale a, implementing equipment type e and varying by time period t. 162
7.3 Process sections contributions to SU and MU processes at scale 2000L on endpoints of (a) human health, (b) ecosystems and (c) resource availability. 173
7.4 Inventory flows percentage contributions in MU process at scale 2000L to midpoints. 174
7.5 Inventory flows percentage contributions in SU process at scale 2000L to midpoints. 175
7.6 Scale-up structures for SU case (a) $t = T_1$, (b) $t = T_2$, (c) $t = T_3$, (d) $t = T_4$, (e) $t = T_5$ and MU case (f) $t = T_1$, (g) $t = T_2$, (h) $t = T_3$, (i) $t = T_4$, (l) $t = T_5$. MU equipment selected (■), SU equipment selected (■) and inactive lines (□) 177
7.7 Scale up KPIs:(a) input demand, (b) cumulative capital investment, (c) total cost, (d) unit cost, (e) total footprint, (f) unit footprint, (g) total GWP emissions, (h) unit GWP emissions . . . 178
7.8 Social vs environmental benefit: (a) quality of life improved vs impact on human health, (b) quality of life improved vs impact on ecosystems. 179
7.9 Uncertainty quantification for representative environmental indicators for different manufacturing scales a and equipment type e: (a) global warming midpoint, (b) freshwater ecotoxicity midpoint, (c) human health endpoint, (d) ecosystems endpoint, (e) resource availability endpoint and (f) total monetised footprint 181

B.1 GSA I-II AdV results 242
B.2 GSA I-II LV results 243
B.3 GSA I-II AAV results 244
B.4 GSA III AdV results 245
B.5 GSA III LV results 246
B.6 GSA III AAV results 247

List of Tables

2.1 Contributions to scheduling in pharmaceutical industries 15
2.2 PSE contributions in investment planning and supply chain design problems 18
2.3 PSE Contributions in optimisation under uncertainty 25
2.4 PSE contributions integrating simulation-optimisation approaches 29

3.1 AdV process recipe. Process information was retrieved from Ferreira et al. (2021), Intelligen, Kis et al. (2022), Merten et al. (2014) 42
3.2 Operational inputs for AdV flowsheet. Ranges are marked as known (Y), if supported by literature, or unknown (N), if ±50% of its nominal value is assumed. Rejection coefficients inputs to obtain the target recoveries are reported in parenthesis. Values retrieved from (a) Ferreira et al. (2021), (b) Jagschies et al. (2018), (c) Fedosyuk et al. (2019), (d) Farnós et al. (2020), (e) Cytiva . . 43
3.3 LV process recipe. Process information was retrieved from Ferreira et al. (2021), Intelligen, Perry and Rayat (2021), Comisel et al. (2021b) 44
3.4 Operational inputs for LV flowsheet. Ranges are marked as known (Y), if supported by literature, or unknown (N), if ±50% of its nominal value is assumed. Rejection coefficients inputs to obtain the target recoveries are reported in parenthesis. Values retrieved from (a) Ferreira et al. (2021), (b) Jagschies et al. (2018), (c) Fedosyuk et al. (2019), (d) Comisel et al. (2021b), (e) Ansorge et al. (2009), (f) Bauler et al. (2020b), (g) Merten et al. (2011), (h) Merten et al. (2014), (i) Cytiva, (j) Perry and Rayat (2021), (k) Intelligen. 45

3.5 AAV process recipe. Process information was retrieved from Ferreira et al. (2021), Intelligen, Cameau et al. (2019), Comisel et al. (2021a), Qu et al. (2015) 46
3.6 Operational inputs for AAV flowsheet. Ranges are marked as known (Y), if supported by literature, or unknown (N), if ±50% of its nominal value is assumed. Rejection coefficients inputs to obtain the target recoveries are reported in parenthesis. Values retrieved from (a) Ferreira et al. (2021), (b) Jagschies et al. (2018), (c) Qu et al. (2015), (d) Thorne et al. (2009), (e) Cecchini et al. (2011), (f) Grieger et al. (2016), (g) Comisel et al. (2021a), (h) Cytiva 47
3.7 Resource costs inputs for AdV, LV and AAV flowsheets. Ranges of inputs are marked as known (Y) or unknown (N). The former implies that the range is supported by literature, the latter implies that range was calculated as ±50% of its nominal value. Values retrieved from (a) Ferreira et al. (2021), (b) Comisel et al. (2021b), (c) Comisel et al. (2021a), (d) Intelligen, (e) Kis et al. (2022) . . . 48

4.1 Computational statistics of the optimisation 86
4.2 Summary of network structures and scales for single-product deterministic case. 3 different demand realisations (low, medium, high) and 3 different process realisations (best-case, median case, worst-case) for LV. 90

5.1 Solution times required to obtain solution for different combinations of $\epsilon_{LV} \times \epsilon_{AdV}$ for the constrained and unconstrained case . . 115

6.1 Computational statistics of the snapshot optimisation framework 145

7.1 Computational statistics of the snapshot optimisation framework 171

A.1 AdV mixed manufacturing inputs (kg/batch) 217
A.2 AdV SU manufacturing inputs (kg/batch) 218
A.3 AdV MU manufacturing inputs (kg/batch) 219
A.4 LV mixed manufacturing inputs (kg/batch) by scale 220
A.5 LV SU manufacturing inputs (kg/batch) by scale 221
A.6 LV MU manufacturing inputs (kg/batch) by scale 222

A.7 AAV mixed manufacturing inputs (kg/batch) by scale 223
A.8 AdV mixed equipment summary by step 224
A.8 AdV mixed equipment summary by step (ctn'd) 225
A.9 LV mixed equipment summary by step 226
A.9 LV mixed equipment summary by step (ctn'd) 227
A.10 AAV mixed equipment summary by step 229
A.10 AAV mixed equipment summary by step (cnt'd) 230
A.11 AdV MU equipment summary by step 231
A.11 AdV MU equipment summary by step (cont'd) 232
A.12 LV MU equipment summary by step 233
A.12 LV MU equipment summary by step (cnt'd) 234
A.13 AdV SU equipment summary by step 236
A.14 AdV SU equipment summary by step (ctn'd) 237
A.15 LV SU equipment summary by step 238
A.16 LV SU equipment summary by step (ctn'd) 239

C.1 Supply chain parameters 249
C.2 Transport distance ρ between upstream node $n \in N^J$ and storage node $n \in N^K$ 249
C.3 Transport distance ρ between storage node $n \in N^K$ and demand zone upstream node $n \in N^Z$ and 249
C.4 Capital costs coefficients c^{CAP} (\$/line) for every section s and scale a for the deterministic cases *best, median, worst* 250
C.5 Facility dependent costs coefficients c^{FAC} (\$/line) for every section s and scale a for the deterministic cases *best, median, worst* . . . 251
C.6 Variable costs coefficients c^{VAR} (\$/batch) for every section s and scale a for the deterministic cases *best, median, worst* 251
C.7 Labour costs coefficients c^{LAB} (\$/batch) for every section s and scale a for the deterministic cases *best, median, worst* 252
C.8 Batch size χ (doses/batch) for each scale a for the deterministic cases *best, median, worst* 252
C.9 Bottleneck time τ (days/batch) for each section s for the deterministic cases *best, median, worst* 252
C.10 Batch time α (days/batch) for each section s for the deterministic cases *best, median, worst* 253

C.11 Batch size χ (doses/batch) for each scale a for each case 254
C.12 Bottleneck time τ (days/batch) for each section s for each case . . 254
C.13 Batch time α (days/batch) for each section s for each case 254
C.14 MU facility cost coefficients c^{FACjl} ($/line) for process section s, product i and scale a. 255
C.15 MU variable cost coefficients c^{LABjl} ($/batch) for process section s, product i and scale a. 256
C.16 MU labour cost coefficients c^{LABjl} ($/batch) for process section s, product i and scale a. 256
C.17 SU facility cost coefficients c^{FACjl} ($/line) for process section s, product i and scale a. 257
C.18 SU labour cost coefficients c^{LABjl} ($/batch) for process section s, product i and scale a. 258
C.19 SU variable cost coefficients c^{LABjl} ($/batch) for process section s, product i and scale a. 258
C.20 Facility cost coefficients c^{FACjl} ($/line) for process section s, equipment type e and scale a. 259
C.21 Labour cost coefficients c^{LABjl} ($/batch) for process section s, equipment type e and scale a. 259
C.22 Variable cost coefficients c^{VARjl} ($/batch) for process section s, equipment type e and scale a. 260
C.23 Normalisation factors μ_f for each endpoint indicator ep 260
C.24 Impacts per batch e^{MP} and e^{EP} for equipment type e process section s and scale $a = 50$ 261
C.25 Impacts per batch e^{MP} and e^{EP} for equipment type e process section s and scale $a = 200$ 262
C.26 Impacts per batch e^{MP} and e^{EP} for equipment type e process section s and scale $a = 1000$ 263
C.27 Impacts per batch e^{MP} and e^{EP} for equipment type e process section s and scale $a = 2000$ 264

D.1 Stainless steel vessels weight data 266
D.2 Stainless steel equipment weight data 266
D.3 50mL to 1L Bags 267
D.4 2L to 3000L Bags 267

D.5 Bag weight estimates 267
D.6 MU 2000L process: key input and output flows 268
D.7 MU 1000L process: key input and output flows 269
D.8 MU 200L process: key input and output flows 270
D.9 MU 50L process: key input and output flows 271
D.10 SU 2000L process: key input and output flows 272
D.11 SU 1000L process: key input and output flows 273
D.12 SU 200L process: key input and output flows 274
D.13 SU 50L process: key input and output flows 275

Chapter 1

Introduction

1.1 Pharmaceutical industrial landscape

The pharmaceutical sector brings value to society through the discovery and deployment of life saving medicines. In recent years, technological advances and enhanced clinical disease understanding have given rise to novel classes of therapeutics based on genetic engineering. In this space, Advanced Therapy Medicinal Products (ATMPs) (EMA, 2020) are an emerging class of therapeutics based on genes, tissues or cells which have demonstrated ground-breaking results in the treatment of cancer, genetic and neurodegenerative diseases as well as acute and chronic conditions.

Specifically, ATMPs for *in vivo* gene therapy rely on introducing a vector carrying a genetic payload in a target patient to fulfill a therapeutic outcome through genetic modifications. *Ex vivo* gene therapy or cell and gene therapy (C&GT) relies on the transplant to the patient of cells that have been previously genetically engineered through the genetic vectors. Vectors can be viral- or non-viral (e.g. mRNA therapeutics) and their genetic payload can be modified based on the application of interest. In this sense, the genetic code can also be tailored to express an antigen protein of interest once administered to patients. As a result of this versatility, genetic-based vectors have paved the way to next-generation mRNA and viral vector-based vaccines (Van Riel and Wit, 2020).

As of 2022, 40 gene therapies resulted approved worldwide, including both viral and non-viral vector-based treatments (Arabi et al., 2022; CHMP, 2019b,a; EMA, 2015, 2016; U.S. Food & Drug Administration (FDA), 2017). The majority of these therapies use viral vectors due to their high transduction efficiency and ability to deliver genetic material effectively (Ghosh et al., 2020; Bulcha et al., 2021), alongside approximately 75% of the 2,000+ recent gene therapy trials employing viral vectors (Yin et al., 2014). Non-viral therapies are emerging alternatives and currently represent a small fraction of approved therapies (Sheridan, 2023).

The global gene therapy market is projected to grow over the coming years. Valued at approximately USD 5.54-9 billion in 2023, it is expected to reach USD 18.20-44.5 billion by 2032, growing at a compound annual growth rate (CAGR) of 18.88-19.4% from 2024 to 2032 (Global Market Insights, 2023; Grand View Research, 2023). Furthermore, a recent study projected that from 2020 to 2034, approximately 1 million patients would be treated through gene therapies,

estimating that tens of thousands of patients would be receiving them annually (Wong et al., 2023). Demand for these products has also risen due to the recent COVID-19 pandemic, with the most prominent vaccine platforms during the recent pandemic being mRNA-based and viral-vector based (CHMP, 2021b,a; Kis et al., 2022; Whitley et al., 2022), enabling the immunisation of the world population with billion of doses delivered (WHO, 2024).

The complexity of genetic-based ATMPs and lack of in-house manufacturing expertise of large multinationals often requires outsourcing manufacturing to smaller contract manufacturing organisations (CMOs) (Capra et al., 2022; Catapult, 2023). These often operate in campaigns to meet mid-term future demands from a set of contracts. In the case of viral vectors, batch production takes from several weeks to months depending on the complexity of the vector and the therapy being produced (Levine et al., 2017; AstraZeneca, 2021). The latter dictates how many doses are manufactured per batch, ranging from 10-1000 doses for targeted gene therapy applications, and million-scale for vaccine production depending on the volumetric process scale (10-2,000L) (Vellinga et al., 2014; Cameau et al., 2019; Comisel et al., 2021b).

In this landscape, genetically engineered pharmaceutical products are paving the way to a paradigm shift in the strategies of process development and planning. Manufacturing processes ought to become product-agnostic, with modifications only at the genetic payload level determining the product application (Kis et al., 2019b). This translates in significant innovations in capacity planning and manufacturing strategies, with an opportunity to flexibly re-purpose multi-product platforms and to shorten process development timelines, allowing for rapid scale-up to meet global growing demands.

Currently, the promise of next-generation flexible pharmaceutical manufacturing networks is hindered by underlying supply chain challenges. CMOs have reported bottlenecks since 2018 (Capra et al., 2022), which highlighted limitations of manufacturing resources at global level. In addition to this, the unforeseen pandemic further exacerbated these issues. With the production of vaccines leveraging on CMO manufacturing capacity, significant backlogs in gene therapy markets were reported due to limited process capacity, uncertain process performance estimates combined with a pressure to urgently scale up for different markets (Capra et al., 2022; Perry and Rayat, 2021; Merten et al., 2011; Ansorge et al., 2009; Bauler et al., 2020a). The case of viral vectors is representative of

this, with critical delays in the delivery of adenoviral vector vaccines during the pandemic (Gardner B, 2020).

With the continued growth of novel biotechnological drugs and pandemic preparedness becoming a growing priority, there is a clear opportunity to develop planning strategies which mitigate the risk of future shortages. Figure 1.1 displays a representation of multi-product production platforms catering for different markets. Specifically, vector-based vaccines are deployed to administration sites and meet clinical demands of patient populations at million-scale. Vector-based *in-vivo* gene therapies target smaller populations and may cater for rare diseases. Demand scale decreases significantly with personalised therapeutics. In this case, viral vectors serve as raw material in cell therapy manufacturing (*ex-vivo* gene therapy), whereby patient-specific batches are manufactured and returned to a single patient (autologous) (Papathanasiou et al., 2020) or manufactured as off-the-shelf products and administered to a compatible population subgroup (allogeneic) (Caldwell et al., 2021). In this landscape, challenges are tied to a need to rapidly scale up for a range of applications. This translates in investment decisions being made before process capabilities are known and optimised (*process* uncertainty) and as a function of future product demands, which are challenging to forecast in terms of timings and scale (*demand* uncertainty). Hence, investment planning becomes *in risk* of not meeting demands once realised with the established capacity.

In addition to the above mentioned challenges in planning under uncertainty, awareness of environmental sustainability of supply chain operations in life sciences and the pharmaceutical sector is becoming increasingly relevant, with recent governmental and industrial initiatives to comply with Net Zero targets by 2030 (HM Government, 2021). Currently, established biopharmaceutical processes often rely on stainless steel equipment, which are water and energy intensive (Budzinski et al., 2022). Recent industrial trends for emerging therapeutics and vaccines undergoing scale-up propose instead a shift towards plastics-based equipment to minimise upfront capital investment and help meet global demands faster, which results in a high degree of solid contaminated waste (Samaras et al., 2022). Hence, there is a growing need to quantify sustainability metrics that help screen decisions and quantify potential trade-offs between cost, environmental and social indicators that steer decision-making.

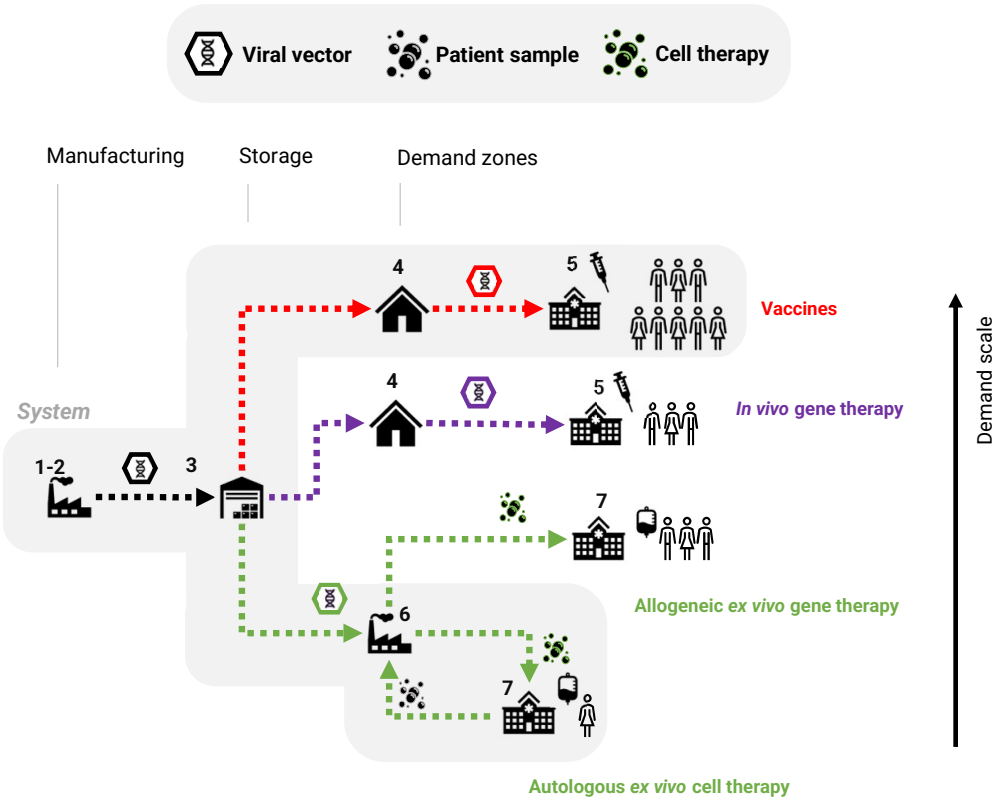


Figure 1.1: General schematic of viral vector supply chains, comprising (1) primary manufacturing, (2) secondary manufacturing/fill-to-finish, (3) centralised storage, (4) decentralised storage, (5) hospitals or administration points as well as (7) specialised leukapheresis sites, (6) cell therapy manufacturing facilities. In this work, we focus on the complexities of viral vector manufacturing for a range of applications and supply chain planning for vaccines and *ex-vivo* gene therapy (*System*).

Through interactions with 40+ industrial partners as part of the Future Targeted Healthcare Hub (EP/P006485/1) it emerges that the adoption of advanced planning tools in biopharmaceutical manufacturing remains limited. This is partly due to the complexity of bioprocesses (Capra et al., 2022), involving numerous dynamics and unpredictable factors (e.g. process performance, demand), which challenges the accurate modelling manufacturing schedules. An additional barrier is the integration of new technologies in existing systems, which rely on manual processes for scheduling and planning by operators, who require further training to operate advanced planning tools. As the industry grows, incentives to overcome these barriers will depend on the ability to quantify adoption margins, demonstrating benefits from reduced production delays, better resource utilisation, reduced costs and footprints.

To this end, computer-aided modelling and optimisation-based tools become *in silico* experimentation platforms to study the impact of various sources of uncertainty on network performance. This helps identify *uncertainty-aware* optimisation solutions which perform well with respect to cost, environmental and social indicators for both predictable and unforeseen uncertainty outcomes, thus reducing the *in risk* component of the investment planning problem.

1.2 Aims & objectives

The research aim of this thesis is the development of a systematic framework to support investment planning and improve supply chain network resilience and sustainability. This is achieved by quantifying comprehensive cost, environmental and social key performance indicators of alternative decisions and the performance of these decisions under uncertainty, which is an indicator of supply chain resilience.

The proposed framework leverages on a quantification and an optimisation step. First, techno-economic modelling and life cycle analysis are performed to quantify performance indicators with respect to cost, footprints and throughputs of alternatives. Second, optimisation models for supply chain planning are developed, which use input data generated in the previous step and help assess decision-making trade-offs.

A key focus of the framework is the integration of uncertainty of input parameters for the optimisation. Quantifiable uncertainties (*known* unknowns) are integrated in the optimisation to seek *uncertainty-aware* solutions. Chapter 3 focuses on the quantification of underlying process-specific manufacturing uncertainties via uncertainty analysis and global sensitivity analysis. In Chapter 4, the generated process parameter distributions are integrated within the optimisation with a comparative assessment of stochastic approaches in two-stage supply chain design problems.

Chapter 5 studies the impact of an *unknown* unknown (rapid increase in vaccine demand) on the performance of network designs obtained through the two-stage stochastic approach. Effectively, we formulate a feasibility problem where we explore the flexibility of a design in meeting two competing demands. The flexibility is called *uncertainty-aware* as it embeds information on the quantified uncertainty in process capabilities in terms of reliability.

Another approach to planning leverages of formulating a multi-stage stochastic problem as shown in Chapter 6. Process uncertainty realises along the time horizon alongside demand uncertainty, which can be quantified based on historical demand data. We investigate the potential of flexible re-configurable manufacturing equipment in rapidly adapting to known uncertainty, which is expected to hold true for unknown unknowns too.

Finally, Chapter 7 proposes the integration of comprehensive sustainability metrics to steer decision-making. Life cycle assessment tools are used to quantify environmental footprints and their respective uncertainty. A deterministic optimisation model is formulated to elucidate cost, environmental and social trade-offs. We conclude with an outlook on integrating uncertainty in impact quantification in the considered investment planning problem.

To conclude, Chapter 8 summarises the contributions of the presented work and the open challenges which motivate further research in the area.

Chapter 2

Literature review

Outputs of this chapter have been disseminated in:

- **Sarkis, M.,** Bernardi, A., Shah, N. and Papathanasiou, M.M. (2021). Decision support tools for next-generation vaccines and advanced therapy medicinal products: present and future, *Current Opinion in Chemical Engineering* 32, 100689.
- **Sarkis, M.**, Bernardi, A., Shah, N. and Papathanasiou, M.M. (2021). Emerging challenges and opportunities in pharmaceutical manufacturing and distribution, *Processes* 9(3), 457.

2.1 Introduction

The pharmaceutical sector is in need of decision-support tools that help meet growing product demands, whilst minimising cost and environmental footprints. Currently, planning tools are not widely adopted by manufacturers in the sector (Marques et al., 2020) (Chapter 1), due to challenges related to lack of staff expertise to operate them, regulatory concerns and, more importantly, uncertainty-related complexities which create discrepancies between pre-planned schedules and the real outcomes. To this end, developing planning tools that are aimed at optimising cost, environmental and social sustainability under various uncertainties and quantifying the benefits of this advanced planning approach would provide incentives to overcome adoption barriers.

Process Systems Engineering (PSE) has extensively researched methodologies and developed frameworks for process planning. Effectively, works differ in terms of systems boundaries, complexities and uncertainties considered. Based on the modelling scope, some recurring challenges become relevant, thus requiring the development of tailored modelling and solution strategies (Reklaitis, 1991; Shah, 1998; Grossmann, 2005, 2012; Barbosa-Póvoa, 2014):

- The modelling challenge: what type of planning models should be developed for the supply chain components to capture key complexities?
- The multi-scale challenge: how to integrate the developed optimisation models across multiple time and space scales?
- The uncertainty challenge: how to account for uncertainty realisations (*known* unknowns) and for *unknown* unknowns in the modelling?
- The computational challenge: how to effectively solve the various models in terms of efficient algorithms?

Extensive contributions in the literature have addressed these problems, generating transferable frameworks and/or tailoring them to specific product applications. We review works in the space of (i) *quantifying* performance metrics for options to be considered within the optimisation and works that have looked at (ii) developing *optimisation*-based tools for planning. We highlight applications in the pharmaceutical sector, transferable tools from other industries, recurring modelling challenges and emerging knowledge gaps.

2.2 Performance quantification

Computer-aided process flowsheeting is a widely established methodology for the quantification of process performance and represents the first step towards the development of tools that can select alternatives based on the decision-making objectives. Techno-economic analysis is a fundamental tool to quantify process performance with respect to throughput and costs. In addition, Life cycle analysis (LCA) is an established methodology for the quantification of environmental impacts throughout products life cycles. In the next sections, we discuss literature that develops computer-aided techno-economic models and LCA to quantify and compare process performance of alternative process configurations, including single-use vs stainless steel equipment, batch vs continuous operating mode and comparisons across scales and products.

2.2.1 Techno-economic modelling

Techno-economic flowsheets rely on an integrated system of mass and heat balance equations, scheduling, equipment sizing, and costing calculations to compute key performance indicators (KPIs) of manufacturing, including batch size, capital and operating cost.

Techno-economic models have been developed to compare batch and continuous manufacturing in the contexts of monoclonal antibodies (mAbs) (Yang et al., 2020; Gupta et al., 2021; Ding and Ierapetritou, 2021) and have highlighted the benefit of continuous processing. Specifically, Bunnak et al. (2016) have argued that cost-effectiveness of continuous manufacturing can be scale-dependent and may be process-specific, thus highlighting the complexity of underlying trade-offs.

Recent contributions have also focused on the emerging sector of Advanced Therapy Medicinal Products (ATMPs), looking to assess techno-economic performance and feasibility of different manufacturing setups. Given that these are processes still under development, uncertainty and sensitivity analyses are valuable tools to quantify the impact of underlying manufacturing uncertainties such as raw material costs and process yields on techno-economic model outputs. Pereira Chilima et al. (2018) presented a cost assessment comparing scale-out (autologous) and scale-up (allogeneic) approaches for cell and gene therapy

(C&GT) manufacturing. A bioprocess and costing model estimates of facility footprint and fixed capital investment of each cell therapy application. Further work by Jenkins and Farid (2018) integrates techno-economic analysis of allogeneic therapies manufacturing with a decisional tool (Jenkins and Farid, 2018). Pereira Chilima et al. (2020) assess the impact of manufacturing decisions on cost of goods (COGs), process robustness, and reimbursement for allogeneic C&GT products. The propagation of uncertainty is assessed via Monte Carlo sampling from pre-defined probability distributions assigned to key process parameters. Local sensitivity analysis (LSA) yields a ranking of impactful parameters including viral transduction and electroporation efficiencies and DSP capacity. Comisel et al. (2021a,b) develop a bioprocess techno-economic model for lentivirus (LV) and adeno-associated vector (AAV) manufacturing with the aim of identifying a cost-effective scale-up strategy from clinical to commercial stages. The authors assess via LSA how raw material cost and process inputs may propagate to COGs/dose for LV-based and AAV-based gene therapy products manufactured via different technologies.

Recent work in techno-economic modelling and sensitivity analysis has also focused on next-generation vaccine platforms, which are largely based on ATMPs. Bioprocess simulations of VV-based vaccines were presented in Ferreira et al. (2021) to conduct assessments between alternative batch and perfusion upstream (USP) setups and assess the impact of USP titer uncertainty on manufacturing KPIs. Kis et al. (2019a) develop SuperPro models for emerging vaccine platforms and integrated a supply chain design optimisation model. LSA is performed to quantify the impact of USP uncertainty on KPIs including total costs and product delivery time. Kis et al. (2022) consider the simultaneous variation of uncertain parameters, including batch failure rate, process scale, production titers, and resource costs. Uncertain inputs are sampled based on Sobol sequences, the techno-economic model is simulated for each sample and outputs are recorded. This helps quantify total production capacities and investments required to meet global demands considering uncertainty for emerging processes. Global Sensitivity Analysis (GSA) is performed to identify the most impactful input parameters, based both on single input and input-input interactions, providing a more comprehensive understanding compared to LSA methods of how input uncertainty propagates to the outputs.

2.2.2 Life cycle assessment

Biopharmaceutical production processes are traditionally water and energy intensive. Stainless steel facilities require extensive use of solvents and buffers in clean-in-place and steam-in-place (CIP/SIP) steps creating significant wastewater streams. Recent industrial trends propose instead a shift towards plastics-based equipment to minimise upfront capital investment and help meet global demands faster. In this context, the use of disposable plastics has caused significant concern with regards to environmental sustainability, due to a high degree of solid contaminated waste (Samaras et al., 2022).

To address this unclear environmental benefit of adopting stainless steel versus single-use (SU) equipment, some LCA studies have focused on quantifying environmental sustainability of these manufacturing alternatives (Budzinski et al., 2022; Pietrzykowski et al., 2013) with a focus on mAbs. Specifically, these works have highlighted that for mAbs manufacturing platforms, SU-based equipment reduces environmental footprint in terms of climate change, human health, ecosystem quality, natural resources and freshwater withdrawal compared to stainless steel. This was linked to the removed requirement for CIP/SIP which reduces use of energy-intensive water for injection (WFI). Additionally, SU in DSP removes the need for caustic chemicals for post-processing treatment and sanitisation, which reduces waste streams.

Bunnak et al. (2016) also use LCA to quantify footprints of continuous mAb production alongside cost of goods (COGs). Perfusion processes are found to consume more energy, water and solid waste compared to batch for the same process throughput. In another study, Neves et al. (2024) evaluate the performance of continuous mAb manufacturing from economic, robustness and environmental perspective, with environmental footprint quantified via water and consumable process mass intensities (PMI). A trade-off between improved water usage and consumables is recorded, with water PMIs reduced in continuous mode as opposed to an increased use of consumables per unit product. This highlights the potential for more comprehensive assessments (e.g. LCA) to unify decision-making.

2.2.3 Remarks

This part of the review highlights that workflows for the quantification of process performance through techno-economic assessments and LCAs are established in the literature and have been used to demonstrate a number of process-specific trade-offs. This creates significant scope for developing optimisation-based decision-support tools that explore alternatives systematically and identify cost-effective and sustainable options, based on resource constraints, product alternatives and target demands.

2.3 Multi-scale modelling

The main outcome of the previously reviewed works for process performance quantification are datasets of throughputs and cost for different process technologies, which may be product specific or process specific. Often, the focus is on manufacturing, however the analysis can be extended to other activities in supply chain nodes.

Based on this techno-economic data, one can formulate optimisation models that may select alternative network configurations and node capacities (*network design* and *capacity planning*) or optimise process schedules (*mid-term planning* and *scheduling*) to meet target demands. The key challenge in doing so is that planning decisions are inherently *multi-scale*, spanning over different time-scales and space-scales. In the next sections, we review contributions in the space of operational planning and strategic planning, with a focus on the modelling assumptions that are relevant to pharmaceutical planning problems.

2.3.1 Scheduling

Scheduling activities are concerned with meeting production requirements including customer orders and production targets imposed by supply chain mid-term and long-term plans. Sequencing tasks over time and assigning suitable resources to enable the fulfillment of these requirements. *Tasks* comprise operations in processing equipment, cleaning, change-overs, transport and *resources* include processing equipment items and utilities as well as operators. Typically, scheduling problems are formulated as follows:

- **Objective**: Fulfilling orders and coordinate activities with reception of raw materials. *Optimising* scheduling operations entails performing task-resource allocations in efficient manner.
- **Inputs**: Target inventory level, task lengths, task yields, unit operating costs, process superstructure.
- **Decisions**: Batch lot sizing, unit task assignments, sequencing of products or individual tasks, timing of individual tasks.
- **Constraints**: Material balances, capacity constraints, resource availability and orders being met in full, non-preemptive processing.

PSE research has focused on the development of transferable and generalised scheduling frameworks and the comparison of approaches with respect to key modelling elements.

Broadly, modelling approaches can be classified as batch-based and recipe-based approaches. The former requires facilities to be represented as sequential processes where batch identity is preserved throughout the planning (Pinto and Grossmann, 1995; Castro et al., 2001; Sundaramoorthy and Maravelias, 2008). The latter approach relies on representing processes as networks characterised by product-specific recipes and process steps. This is well-suited for multi-purpose facilities where the complexity of resource allocations, task timings, material flows is best represented as a network, specifically as states-task networks (STN) and resource-task networks (RTN) (Kondili et al., 1993; Shah et al., 1993; Pantelides, 1994).

Within each modelling approach a number of studies have addressed problems related to time-representation (Castro et al., 2009; Schilling and Pantelides, 1996; Ierapetritou and Floudas, 1998). On the one hand one could represent models via precedence-based constraints and task-task sequencing constraints. On the other hand, the complexity of certain processes may require monitoring of material balances over a time-grid, described as a discrete or continuous time set.

These fundamental works form the back-bone of the multi-stage representations or network-representations in pharmaceutical scheduling frameworks. Recent contributions in this space are summarised in Table 2.1 which highlights aims of the works and the system type of modelling implemented.

Table 2.1: Contributions to scheduling in pharmaceutical industries

Research area	Authors	Industry	Multi-site	Modelling	Sustainability?
Scheduling	Stefansson et al. (2006, 2011)	Pharmaceutical	Single	Batch-based	✗
	Moniz et al. (2012, 2014)	Pharmaceutical	Single	Recipe-based	✗
	Lakhdar et al. (2005, 2007)	Pharmaceutical	Single	Bioatch-based	✗
	Liu et al. (2014)	Biopharmaceutical	Single	Batch-based	✗
	Kabra et al. (2013)	Pharmaceutical	Single	Batch-based	✗
	Vieira et al. (2016, 2017, 2019)	Biopharmaceutical	Single	Recipe-based	✗
Planning	Siganporia et al. (2014)	Pharmaceutical	Multi	Batch-based	✗
	Oyebolu et al. (2017, 2019)	Pharmaceutical	Multi	Batch-based	✗
	Amaro and Barbosa-Póvoa (2008)	Pharmaceutical	Multi	Recipe-based	✗
	Sousa et al. (2011)	Pharmaceutical	Multi	Batch-based	✗
	Susarla and Karimi (2012)	Pharmaceutical	Multi	Batch-based	✗
	Georgiadis and Georgiadis (2021)	Biopharmaceutical	Multi	Node rate	✗
	Ibrahim et al. (2021)	Biopharmaceutical	Multi	Node rate	✗
	Badejo and Ierapetritou (2024)	Pharmaceutical	Multi	Node rate	✓

2.3.1.1 Single-site planning

In pharmaceutical contexts, primary manufacturing is often operated in campaign mode which is best represented by the integration of a mid-term campaign planning level. This was first proposed in Papageorgiou and Pantelides (1996), to augment network-based short-term scheduling optimisation. For instance, Stefansson et al. (2006) integrate campaign and scheduling planning assuming a multi-stage process with a hierarchical formulation. The upper campaign level determines a plan to fulfill long-term orders and raw material procurement, a lower level performs unit-task assignments in parallel equipment. The work is further refined in Stefansson et al. (2011) with an industrial-based case study and highlights superior computational performance of continuous time formulations compared to discrete-time counterparts.

A multi-stage representation of manufacturing steps is suitable for biopharmaceuticals applications as well. Manufacturing in the sector complies with stringent regulations on contamination and it is reasonable to assume that batch identity is preserved throughout manufacturing. Hence, Lakhdar et al. (2005) propose a model where detailed modelling of unit-scheduling operations is not considered and capacity is represented as USP and downstream (DSP) suites.

This mixed-integer linear program (MILP) is augmented by Liu et al. (2014) who consider maintenance planning and consider decay in performance of chromatographic resins.

The campaign planning problem can also be integrated with additional activities, including equipment redesign (Moniz et al., 2012) and non-regular production (Moniz et al., 2014). In these works, models rely on RTN MILP formulations to handle the complexities and time description of the retrofitting activities and non-regular scheduling in pharmaceutical manufacturing. Biopharmaceutical planning problems based on STNs and RTNs are presented in Kabra et al. (2013) and Vieira et al. (2016, 2017) respectively. The latter works integrate performance decay considerations of downstream resins as shown in Liu et al. (2014), quantifying trade-offs between profit maximisation and maintenance downtime. Ultimately, the authors present a comprehensive planning tool for industrial users in Vieira et al. (2019).

2.3.1.2 Multi-site planning

The previous works focus on planning within single-sites of the supply chain and capture mid-term and short-term activities. Effectively, pharmaceutical manufacturers may orchestrate production amongst different facilities, which motivates research in multi-site planning. Expanding space system boundaries leads to increased problem dimensionality and thus requires simplifications or tailored solution strategies. Lakhdar et al. (2007) consider a long-term planning problem for multiple facilities. Despite an increased problem size compared to the single-site problem Lakhdar et al. (2005), the formulation remains tractable as it ignored detailed unit-level scheduling.

This detail is instead preserved in the lower-level of the RTN formulation by Amaro and Barbosa-Póvoa (2008) which is integrated with a upper-level RTN formulation based on *macro*tasks and *macro*states. The multi-level formulation preserves short-term scheduling detail and achieves tractability of the supply chain planning problem.

Similarly, Sousa et al. (2011) and Susarla and Karimi (2012) address global supply chain planning problems for multinational pharmaceutical industries. The dimensionality of the resulting multi-site problem requires the development of tailored solution strategies. Sousa et al. (2011) propose a spatial and temporal

decomposition scheme to tackle the computational effort, whereas Susarla and Karimi (2012) propose a set of linearisation procedures that are suited for the structure of the formulation.

A large multi-site planning problem is also addressed by Badejo and Ierapetritou (2024). The authors model manufacturing nodes as single-stage steps and aggregate production levels are determined by the optimisation disregarding detailed scheduling information. This is also done by Ibrahim et al. (2021) and Georgiadis and Georgiadis (2021), looking at integrated planning of COVID-19 vaccination campaigns and manufacturing networks and focusing on providing breakdowns of alternative campaign solutions to steer decision-making.

2.3.2 Capacity planning

The single-site and multi-site operational planning problem presented in the previous section rely on a set of utilisation and resource constraints which are imposed by a given manufacturing design or network structure. The capacity planning problem introduces a new degree of freedom which is the possibility of resource expansion/reduction. As lead times brought by construction/decommissioning, have duration of months, years with respect to short-term operations the capacity planning problem is appropriately described over a longer time horizon.

- **Objective**: The goal of capacity planning is to ensure availability of enough production capacity to meet projected demands of a company's product portfolio. *Optimising* capacity plans involves identifying cost-effective investment plans, minimising costs and maximising utilisation.
- **Inputs**: Target demand to fulfill, scale-dependent capacity and costs, environmental footprints, network superstructure.
- **Decisions**: Investment in new capacity throughout the time horizon or decomissioning existing capacity, node activation, transport links established.
- **Constraints**: Material balances, logic constraints imposing capacity availability only after the required construction and lead times, maximum allowable capacity expansions, sales demand constraints.

Case studies in the pharmaceutical industry often require capacity representations that mirror multi-product plant operations. Key publications in this space are highlighted in Table 2.2.

2.3.2.1 Capacity expansion

Table 2.2: PSE contributions in investment planning and supply chain design problems

Research area	Authors	Industry	Multi-site	Batch-based	Sustainability?
Scale-up	Rotstein et al. (1999)	Pharmaceutical	Single	Batch-based	✗
	Papageorgiou et al. (2001)	Pharmaceutical	Single	Batch-based	✗
	Levis and Papageorgiou (2004)	Pharmaceutical	Single	Batch-based	✗
	Tsang et al. (2007a,b)	Biopharmaceutical	Single	Batch-based	✗
	Sundaramoorthy et al. (2012b)	Pharmaceutical	Single	Batch-based	✗
	Siganporia et al. (2014)	Biopharmaceutical	Single	Batch-based	✗
Network design	Kis et al. (2019a)	Pharmaceutical	Multi	Batch-based	✗
	Triantafyllou et al. (2022)	ATMPs	Multi	Batch-based	✗
	Bernardi et al. (2021)	ATMPs	Multi	Batch-based	✗
	Tsiakis et al. (2001)	Process	Multi	Batch-based	✗
	Tsiakis and Papageorgiou (2008)	Chemical	Multi	Batch-based	✗
	Sousa et al. (2008)	Food	Multi	Batch-based	✗
	Laínez et al. (2009)	Food	Multi	Batch-based	✗
	Liu and Papageorgiou (2013)	Agro-chemical	Multi	Batch-based	✓
	Liu et al. (2020)	Process	Multi	Node rate	✗
	Hugo and Pistikopoulos (2005)	Process	Multi	Node rate	✓
	Guillén-Gosálbez et al. (2010)	Chemical	Multi	Node rate	✓
	Elias Mota et al. (2020)	Paper	Multi	Node rate	✓
	Duarte et al. (2022)	Vaccines	Multi	Node rate	✓

Rotstein et al. (1999) model production capacity for biopharmaceuticals as a set of suites, each comprising an upstream and downstream line (Figure 2.1c). Further work by Papageorgiou et al. (2001) is based on the same configuration and capacity expansion decisions are made throughout the time horizons, considering lead times associated scale-up, start-up time and qualification runs of each suite. Demands for future products entering the portfolio are assumed to be known. The model determines suite expansions and production levels across multiple sites and within each time period a length of production campaign in each suite is selected. The authors also consider the company's trading structure

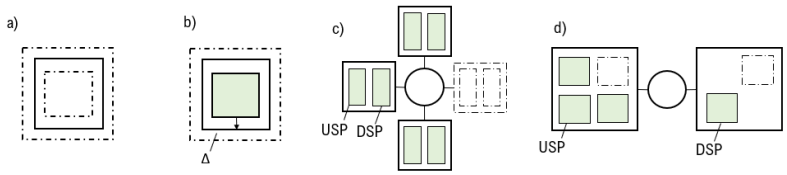


Figure 2.1: System representation in capacity planning problems. General approaches: (a) discrete set of available sizes, (b) capacity increments (Δ). Approaches based on pharmaceutical industry facilities: (c) parallel suites comprising a production and purification line (Rotstein et al., 1999; Papageorgiou et al., 2001), (d) decoupled upstream and downstream suites (Gatica et al., 2003a).

in the modelling. In Siganporia et al. (2014), the bioprocess capacity representation is equivalent to Papageorgiou et al. (2001). Investing into a new facility translates in a new upstream and downstream suite available to be retrofitted according to the selected operating mode.

The manufacturing capacity model presented in Gatica et al. (2003a) and Tsang et al. (2007a,b) differ from the above, as they decouple upstream and downstream suites respectively (Figure 2.1d) and focus on a single-site. The amount of product produced in each time period is defined by campaign lengths. Products are allocated to suites as they become available throughout the time horizon and in relation to their demand. Capacity expansions entail investment into USP and DSP suites accordingly, which become available after defined construction lead-times.

Alternatively, Sundaramoorthy et al. (2012a) present a capacity planning framework for an increased number of pharmaceutical plants. The capacity expansion decisions are modelled from a set of discrete sizes (Figure 2.1a), which mirrors the rationale of scale-selection which may occur throughout the phases of scale-up/scale-out.

Capacity increments can be defined more generically without accounting for a multi-suite facility description or scale-selection. For instance, Laínez et al. (2009) address the capacity planning problem for flexible polymer supply chain networks with a description of capacity expansion decisions as discrete increments in the production rates (Figure 2.1b). Manufacturing and distribution tasks are either batch or continuous and an STN-MILP for the entirety of the supply chain schedules and plans capacity expansions. This incremental capacity is also proposed by Liu and Papageorgiou (2013) which also formulate capacity

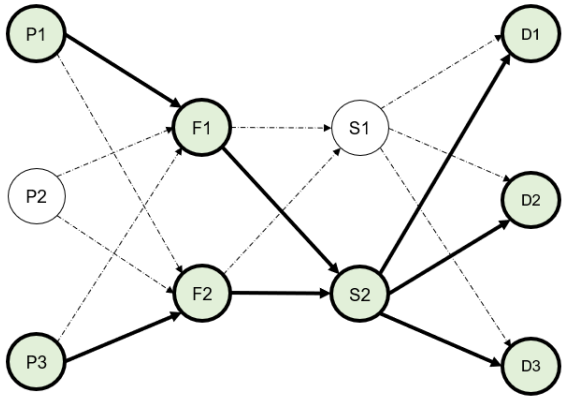


Figure 2.2: Network design: active nodes within P_N, F_N, S_N, D_N allow links and shipments

expansion decisions as increments in the bounds on the attainable production capabilities in their MILP formulation. A proportional capacity expansion entails a maximum capacity increment that is proportional to the initial size of the facility, whereas a cumulative capacity expansion allows a maximum increment to pre-defined value, regardless of original facility sizes.

2.3.2.2 Network design

The supply chain *design* problem requires the introduction of facility location and structural links as decision variable a within the optimisation. If a node activates, inbound and outbound material flows are allowed (Figure 2.2). This type of modelling adds the geographical component in the optimisation problem and translates into a possibility of comparing centralised and distributed supply chain structures.

Recent contributions in the space of pharmaceuticals by Kis et al. (2019a) have highlighted how centralised vs decentralised networks perform in the context of responsive large vaccine deployment. This directly addresses emerging questions regarding the quantification of benefits of distributed manufacturing. With a focus on next-generation vaccine technologies (e.g. RNA-based vaccines, virus-like particles). The sensitivity of the optimisation solutions and calculated costs to product titer uncertainty is assessed, quantified via techno-economic modelling. The exercise reveals intensified processes (e.g. laminar flow cabinet manufacturing) can render small-scale distributed manufacturing economically viable compared to centralised manufacturing, highlighting the need to consider

future bioprocess capabilities for a comprehensive assessment.

Distributed supply chain networks are increasingly relevant in personalised healthcare, where manufacturing closer to demand sites reduces return times. Triantafyllou et al. (2022) present a supply chain planning problem for chimeric antigen-receptor (CAR) T-cell therapy clinical trials. The model captures the patient specificity of the manufactured batches and considers a constraint on the return time of the manufactured therapies to patients. Thus, it requires detailed day-to-day scheduling constraints which lead to a computationally expensive problems making data-driven approaches to solve the optimisation necessary (Triantafyllou and Papathanasiou, 2024). The formulation model is also further developed in Bernardi et al. (2021) to account for on-site storage at the cell-therapy manufacturing facility, which proves to be an efficient strategy to improve the utilisation and scheduling of the parallel lines.

Overall, supply chain design has been largely addressed for other sectors and has its fundamentals on seminal contributions, including the work by Tsiakis et al. (2001). This represent a general framework for supply chain design, where an MILP model is developed to decide on locations and links between facilities, comprising warehouse, distribution centers and customer zones allocation. The model is formulated in terms of time-averaged quantities (i.e. steady-state), therefore the time-evolution of the structure is not considered, with production levels are calculated as a continuous rate constrained by an upper and lower bound. Tsiakis and Papageorgiou (2008) augment the formulation with the possibility to outsource production if demand cannot be met by the installed capacity.

Multi-scale activities can also be modelled through a multi-level integration of long-term planning and detailed scheduling models. In Sousa et al. (2008), an aggregated level determines the supply chain structure and the design decisions are fixed at the lower level. Production plans are determined in the upper level and the detailed operational lower model is used to test the accuracy of such targets. Feedback from the lower level is introduced with an analytical methodology that imposes constraints on the upper level and successfully prevents bottlenecks and infeasible targets. In Liu et al. (2020) the objective is to identify low-cost supply chain structures for new product formulations in the fast-moving consumer goods industry, assessing different network structures. An aggregate static model determines locations of plants, warehouses and distribution centers, whereas a

dynamic model determines cost-effective production and distribution plans. The selection of facility sizes considers small, medium and large capacity, which is characterised by a size-dependent rate of production (Figure 2.1a).

In the above works, candidate facilities locations belong to a pre-defined set. Alternative approaches considering supply chain design for the energy sector (Almansoori and Shah, 2006, 2009; Panteli et al., 2018) represent candidate locations as part of a grid. These works determine a number of facilities to be installed in each grid, without modelling each facility specifically; if facility-specific operations are to be modelled, a grid approach would lead to an increased number of decision variables indexed by grid number.

2.3.3 Remarks

The choice of representing processes as multi-stage processes or networks depends on material handling restrictions and the application that is being studied. Overall, network-represented processes tend to lead to more compact formulations, however they require additional constraints to preserve batch-identity and mimic industrial operations. In this sense, multi-stage representation is well-suited for biopharmaceutical applications, with capacity represented as suites of process sections mostly operated in campaigns.

Currently the literature does not differentiate the use of single-use (SU) and multi-use (MU) equipment, which have different set-up times and flexibility in switching scales. From a modelling perspective, this is feature worth adding to capacity planning frameworks.

Overall, by adopting this suite-based modelling strategy, one simplifies the formulation and avoids embedding scheduling detail in capacity planning problems for pharmaceuticals, which should reduce computational effort. However, upon model development it is worth testing this assumption, by assessing whether variations in the expected performance in short-term activities impact the effectiveness of strategic plans. Hence, it is worth focusing on whether this has been addressed in frameworks for optimisation under uncertainty.

2.4 Planning under uncertainty

The previous part of the literature review has highlighted that the backbone of modelling is well-established with a recurring trade-off between modelling detail and computational tractability. In the bio-pharmaceutical sectors and ATMPs, manufacturers deal with a number of uncertainties simultaneously, often manifesting asynchronously and across different scales and thus requiring a quantification of *resilience* of network design and plans to this multivariate uncertainty.

In the following sections, we review works in the space of planning under uncertainty across different scales, with a focus on understanding assumed uncertainty time-scales, underlying assumptions of probability distributions and optimisation strategies. Some approaches have focused on optimising given some information of uncertainty (Section 2.4.1), whereas other execute sequential or integrated simulation routines (Section 2.4.2).

2.4.1 Optimisation under uncertainty

Optimisation approaches for planning under uncertainty are aimed at identifying solutions that perform well under a range of uncertainty realisations. A number of approaches exist to obtain such solutions, the selection of which depends on the information available regarding the uncertainty and the recourse actions available upon uncertainty realisation. This corresponds to a moment of time where the uncertainty is *measured* by the system. If a recourse action is available, then one can obtain uncertainty realisation (i.e. scenario)-dependent solutions. Alternatively one may seek a solution which is feasible for all realisations of uncertainty, especially if no recourse actions are available. Broadly, strategies for optimisation under uncertainty include:

- Stochastic programming: this approach models uncertain parameter realisation through discrete probability distributions and leverages on scenario trees. The optimisation is formulated with an objective function which takes the expectation over a number of scenarios (Figure 2.3).
- Robust optimisation: the optimisation seeks solutions that are feasible for all realisations of uncertainty parameters. This is suitable when there are

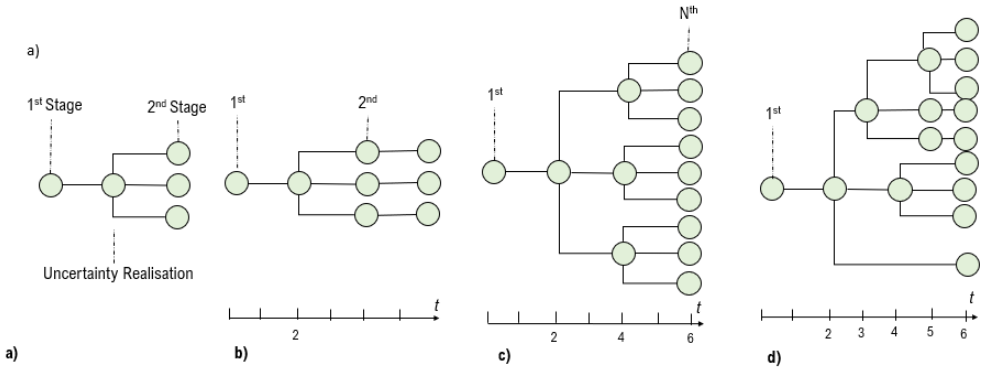


Figure 2.3: Problem representations for optimisation under uncertainty, specifically stochastic programming and robust optimisation: (a) Two-stage, (b) two-stage multi-period, (c) multi-stage multi-period with exogenous uncertainty (with realisations at $t = 2, 4, 6$), (d) multi-stage multi-period with endogenous uncertainty (with realisations timings depending on the decisions made at each stage).

no recourse actions available and one wants to minimise the violation in case an uncertain value is observed.

- Chance-constraint programming: this approach allows constraints to be satisfied with a probability *a priori*, individually or jointly (Charnes and Cooper, 1959). If a scenario-based approximation is proposed, one could see this approach as a solution of a stochastic program with probabilistic constraints.

2.4.1.1 Short-term uncertainty

Depending on the purpose of the modelling different approaches have been proposed to model short-term uncertainty and decision-making under uncertainty in scheduling contexts.

If no recourse actions are available upon realisation of uncertainty, robust optimisation of schedules is a suitable approach. This is highlighted in Janak et al. (2007), where uncertain parameters are described by a known probability distribution function. The problem is reformulated as its deterministic robust counterpart. Similarly, Letsios et al. (2021) approach an industrial resource allocation problem through a robust scheduling approach where first-stage decisions are determined with exact lexicographic scheduling and second-stage decisions are derived through approximate rescheduling.

Table 2.3: PSE Contributions in optimisation under uncertainty

Research area	**Authors**	**Strategy**	**Recourse**	Uncertainty
Scheduling	Balasubramanian and Grossmann (2004)	Stochastic	✓	Demand
	Bonfill et al. (2005)	Stochastic	✓	Demand
	Wu and Ierapetritou (2007)	Stochastic	✓	Demand
	Lin et al. (2004)	Stochastic	✓	Demand
	Schilling and Pantelides (1997)	Robust	✗	Demand
	Janak et al. (2006)	Robust	✗	Demand
	Schmidt and Grossmann (1996)	Stochastic	✓	Demand
	Colvin and Maravelias (2008)	Stochastic	✓	Demand
	Lakhdar et al. (2006)	Chance-constraint	✗	Process
Planning	Gupta and Maranas (2000)	Stochastic	✓	Demand
	Amaro and Barbosa-Póvoa (2009)	Stochastic	✓	Demand
Capacity planning	Gatica et al. (2003a)	Stochastic	✓	Demand
	Levis and Papageorgiou (2004)	Stochastic	✓	Demand
	Tsang et al. (2007a)	Stochastic	✓	Demand
	Sundaramoorthy et al. (2012a)	Stochastic	✓	Demand
	Maravelias and Grossmann (2001)	Stochastic	✓	Demand
	Salema et al. (2007)	Stochastic	✓	Demand
	Cardoso et al. (2013)	Stochastic	✓	Demand
	Mestre et al. (2015)	Stochastic	✓	Demand

Alternatively, distributional information may not be available in real-industrial scenarios for instance because of lack of measurements. Wiebe et al. (2019) explore the use of robust optimisation in non-linear pooling problems considering uncertainty input concentrations in petrochemical refining contexts.

Traditionally robust optimisation results in overly conservative solutions. A multi-stage adaptive robust optimisation (ARO) helps reduce the conservativeness of the solution, as shown in Lin et al. (2004) and Janak et al. (2007). This approach allows the consideration of many uncertain parameters include processing times of tasks, market demands prices of products and materials. In the first stage before uncertainty realisation, a robust optimisation is solved, with the uncertainty set adapting through the progression of stages.

Effectively, in Lakhdar et al. (2006) the optimisation under uncertainty also

considers no recourse actions. The authors focus on campaign planning and the problem of fermentation titer uncertainty, which is demonstrated to be a critical manufacturing uncertainty through sensitivity analysis. The proposed formulation leverages on a set of chance-constraints which are reformulated based on the P-percentile value of probability distribution for each product. This allows the development of campaign plans which are robust without any recourse action to uncertainty in fermentation titers to a probability level of P.

If recourse actions and distribution information are available, stochastic programming is often the chosen alternative. Balasubramanian and Grossmann (2004) address demand uncertainty with a two-stage programming approach. Processing mode selection, sequencing and batch sizing as first stage decisions common to all scenarios. Sales decisions and unsatisfied demand are determined in the second-stage. Bonfill et al. (2005) instead implement stochastic programming techniques to address uncertainty in processing times during scheduling. First-stage decisions determine production sequences, unit-task assignment and whereas timings of the schedules are determined in a second stage. Similarly, Wu and Ierapetritou (2007) propose the simultaneous solution of production planning and scheduling and considers uncertainty with a three-stage stochastic approach.

Stochastic programming approaches have found important application in clinical trial planning. The work by Maravelias and Grossmann (2001) forms the basis of the clinical trials planning framework proposed in Colvin and Maravelias (2008). Uncertainty is modelled as as *endogenous*, which complicates the scenario tree as it no longer fixed *a priori*. The integration of this framework with resource-constrained and risk-constrained expansion and outsourcing decisions was presented in Colvin and Maravelias (2009).

In mid-term planning contexts for multi-site systems, Gupta and Maranas (2000) and Gupta et al. (2003) formulate a two-stage stochastic supply chain planning model. Production decisions are made before the realisation of the demands. Recourse decisions are related to distribution plans, including inventory levels and transportation flows. Similarly, Amaro and Barbosa-Póvoa (2009) address the problem of pharmaceutical supply chain planning under demand uncertainty. A two-stage stochastic optimisation is applied; however, the detailed scheduling lower level is removed from the formulation in Amaro and Barbosa-Póvoa (2008) to preserve computational tractability.

2.4.1.2 Long-term uncertainty

The time-scale of recourse actions is suitable for formulating stochastic problems in mid-term planning contexts, if information on probability distributions is available. This holds for long-term planning too and the literature has seen a large reliance on stochastic programming for investment planning under uncertainty.

Most of the capacity planning frameworks described in Section 2.3.2 were formulated to tackle the problem of investment planning under clinical trial/demand uncertainty. Rotstein et al. (1999) formulate a two-stage stochastic programming problem considering clinical trial outcomes uncertainty. First stage variables include initial capacity investments, product selection and initial allocation of manufacturing resources to products. Second stage recourse decisions consist in additional capacity investment or sale of capacity, re-allocation of products to resources. An iterative solution procedure eliminates scenarios with low probability and reduces the computational effort.

Gatica et al. (2003a) extend the work by Rotstein et al. (1999) to generate capacity plans under uncertainty over a long time horizon, with a multi-stage stochastic program. At each stage the demand of each product is updated according to four different scenarios, including high, target, low success or failure of trials. The formulation maximises expected NPV and risk mitigation is introduced through a constraint on the expected downside risk. With a similar rationale, Tsang et al. (2007a,b) focus on capacity planning under uncertainty for multiple vaccines and extend the work by Gatica et al. (2003a) to evaluate different risk metrics (expected downside risk, opportunity value, value-at-risk and conditional value-at-risk).

A series of works have demonstrated the suitability of stochastic programming to solve location (first-stage) problems and allocations (second-stage). These include contributions from Salema et al. (2007) and Cardoso et al. (2013) consider network structures with integration of reverse logistics activities. This entails that the optimisation chooses sizing, location of plants and distribution nodes and establishes forward and reverse flows. Similarly, Mestre et al. (2015) develop a framework for hospital networks design and discuss two variable partition strategies for the location-allocation decisions.

Stochastic problems rapidly increase in size due to either modelling detail or

scenario tree size. By modelling for both detailed scheduling and capacity planning under clinical trial uncertainty, Maravelias and Grossmann (2001) requires a decomposition approach or heuristic decomposition approaches (Maravelias and Grossmann, 2004) for the solution. When combining efforts by Papageorgiou et al. (2001) on network trading structure modelling and Gatica et al. (2003a) into a single framework, Levis and Papageorgiou (2004) develop a hierarchical decomposition algorithm, where an aggregated MILP is solved first to determine strategic decisions and fix them in the detailed level to reduce the variable space. Efficient solution strategies are sought by Rotstein et al. (1999) and Gatica et al. (2003b) themselves, presenting scenario-reduction and aggregation methods and discussing the risk of loss of accuracy. Similarly Sundaramoorthy et al. (2012a) focus on capacity planning under uncertainty for a multi-site network relying on stochastic programming and propose a decomposition scheme in Sundaramoorthy et al. (2012b) to address problem size.

2.4.2 Simulation-optimisation

Decision-making frameworks can embed simulations through different approaches. Key contributions in the PSE literature are summarised in Table 2.4. First, it is worth differentiating *static* simulation or *dynamic* simulation as follows.

- Static simulation: this approach is used to represent systems at a specific point in time without considering changes over time. These models assume that the system being studied remains constant and does not evolve or change over time.
- Dynamic simulation: this approach is used to represent systems where the behaviour changes over time. These models take into account how variables evolve and interact with each other over time, allowing for a comprehensive understanding of dynamic behaviour of the system.

2.4.2.1 Static simulation

A static simulation may help assess the propagation of uncertainty for a certain optimisation solution and helps pinpoint the performance under uncertainty of the solution. In this context, Monte Carlo (MC) and quasi-Monte Carlo (QMC)

Table 2.4: PSE contributions integrating simulation-optimisation approaches

Research area	Authors	Decision rule	Simulation	Uncertainty
Reactive scheduling	Janak et al. (2006)	Robust	Discrete	Demand
	Ryu and Pistikopoulos (2007)	mp	Discrete	Demand
	Li and Ierapetritou (2010)	mp	Discrete	Demand
Sim-Opt	Gökalp and Branke (2020)	Heuristics	Discrete	Demand, disruptions
	Subramanian et al. (2001)	Det. Optimisation	Discrete	Process, disruptions
	Oyebolu et al. (2017)	Heuristics	Discrete	Process
Rolling horizon	Bose and Pekny (2000)	Optimal control	Dynamic	Demand
	Perea-López et al. (2001)	Optimal control	Dynamic	Demand
	Subramanian et al. (2013)	MILP, Control	Dynamic	Demand
	Aguirre et al. (2018)	MILP, Control	Dynamic	Demand
	Badejo and Ierapetritou (2022)	Rolling horizon	Discrete	Demand
	Badejo and Ierapetritou (2024)	Rolling horizon	Discrete	Demand
MC simulation	Gjerdrum et al. (2001)	Agent-based	Static	Process
	Hung et al. (2006)	Agent-based	Static	Process
	Marques et al. (2017)	MILP	Static	Demand, trial outcomes

methods find important application in evaluating uncertainty propagation to model or optimisation model outputs.

Some works integrate a MC simulation methodology prior to the optimisation. For instance, Sample Average Approximation (SAA) schemes (Santoso et al., 2005) have been proposed to tackle a large-scale supply chain design problem and rely on MC simulations to calculate the expected value of the second stage decisions. A sample of the scenarios which approximate the entire set is obtained and this provides an approximated expected objective value. In this case, simulation is used to generate scenarios for the stochastic optimisation.

Conversely, the frameworks for stochastic simulation presented in Gjerdrum et al. (2001) and Hung et al. (2006) rely on MC simulations to generate confidence bounds on the service levels of the supply chains. These are effectively used to compare supply chain policies and compute probabilities of stock-outs, thus informing decision-makers on the risk of implementing either strategies and go back to the policy design stage. The propagation of uncertainty is in itself static: uncertainty samples are generated in an outer iteration loop for each sample and the supply evolution over time is simulated considering agent interactions.

Along this rationale, MC simulation can be used to quantify solution performance and identify critical input parameters, effectively conducting a sensitivity analysis. In Lakhdar et al. (2006), uncertainty of fermentation titre in biopharmaceuticals results a single critical parameter worth mitigating against (i.e. batch size), disregarding potential correlations with other process parameters and enabling the formulation of a deterministic chance constraint.

In Marques et al. (2017), a two-step approach is proposed to address the clinical trial planning problem. First, MC sampling is used to generate a large number of instances of product demand and clinical trial outcomes. An MILP is solved for each instance and the solution is recorded, thus obtaining a probability density function for the objective function and probabilistic occurrence of decision variables. This can be repeated to compare different process design configurations and capacity planning decisions set *a priori*.

2.4.2.2 Rolling-horizon optimisation

It is often the case in the literature that an optimisation routine is simulated *dynamically* through a rolling horizon approach, with the model solved repeatedly and the planning interval moved forward in time during each solution step. In some instances, this is done to reduce the computational effort in the solution, in others, this is done to mimic the real decision-making practice.

For instance, reactive scheduling frameworks have been proposed to adapt the existing schedule to subverted system conditions. In Janak et al. (2006) the authors address short-term unforeseen events, such as unit shutdowns and addition or modification of orders. An algorithm is proposed to avoid full rescheduling at each iteration. As unforeseen orders arrive, the STN-MILP formulation determines which tasks can be fixed as unaffected by uncertainty, thus allowing the rest of the horizon to be rescheduled with reasonable computational effort.

This reported computational expense for online rescheduling can cause delays upon real-plant implementation. Hence, Li and Ierapetritou (2008) directly address the response time of these techniques, by formulating a multi-parametric (mp) programming problem. This approach is derived from the principles of mp-MPC, which looks solving all possible outcomes of uncertainty *a priori*, thus obtaining look up maps for online implementation. Further studies in this space include the work by Ryu et al. (2007), with the development a proactive

rescheduling framework relying on mp-MILP, where rescheduling is performed via simple function evaluations without further optimisation. Similarly, Li and Ierapetritou (2010) address the integrated mid-term production planning and scheduling optimisation problem incorporating of valid capacity information into the planning level. The key challenges lie in the computational expense required to obtain the look-up maps in the first place, which may limit the scalability of the mp approach.

Rolling horizon solutions of mixed integer optimisation and formal MPC find extensive application in supply chain planning under uncertainty. Bose and Pekny (2000) develop an MPC framework where the forecasting model calculates the target inventory for future time periods and the optimisation model yields the scheduling tasks necessary to achieve these targets. Perea-López et al. (2003) employ MPC supply chain management and compare centralised and decentralised management approaches. Flows of material and information (orders) in this case are modelled as difference equations. The supply chain planning model in Liu et al. (2012) uses an MILP optimisation to determine production sequences and MPC applied to manipulate production levels and maintain inventory and price at desired levels as uncertainties in price and demand realise. Subramanian et al. (2013) use MPC in supply chain inventory management and demonstrate the possibility to integrate scheduling and control a multi-product manufacturing facility. Aguirre et al. (2018) propose hybrid decomposition approaches to be coupled with MPC to tackle large-scale problem sizes in planning.

The previously discussed capacity planning work by Siganporia et al. (2014) uses rolling horizon approaches to solve the optimisation to tackle tractability and effectively simulate real-world planning activities. The demand prediction horizon would span n months ahead and the planning window would shift progressively. Similarly, this is the approach proposed by in the clinical trial planning work by Colvin and Maravelias (2008) and supply chain planning frameworks presented in Badejo and Ierapetritou (2022) and Badejo and Ierapetritou (2024). Here disruptions and demand are considered *disturbances*, realised at the optimisation level only once as the prediction horizon is shifting.

2.4.2.3 Simulation-based heuristics

Alongside mathematical programming approaches, a few works have focused on exploring the potential of combining simulation and optimisation modules to broaden the capabilities of planning tools in handling multiple sources of uncertainties simultaneously. In Subramanian et al. (2001), the authors argue that single-level mathematical programs have inherent computational limitations in capturing simultaneous stochastic elements.

To this end, the works of Subramanian et al. (2001), Zapata et al. (2011) and Chen et al. (2012) present the development of discrete-event simulation-optimisation (Sim-Opt) frameworks which combine a stochastic simulation block and an optimisation block. The latter is used to substitute the typical local and myopic rules in discrete-event simulation methods to resolve conflicts and perform decisions for different simulated uncertainty outcomes. Upon iterative simulation, statistical trends can be tracked and answers to questions related to corporate policies can be obtained.

This stream of research has continued in recent years. Oyebolu et al. (2019) consider that continuous bio-pharmaceutical processes are more prone to failure. Probability of disruption is dependent on process run-time, which is determined by production schedules. A discrete-event simulation generating instances of equipment disruption and demands is integrated with a scheduling environment based on heuristics based policies. With a similar rationale, Gökalp and Branke (2020) address the pharmaceutical R&D pipeline management problem under clinical trial outcome and duration uncertainty and substitute the heuristics with an Approximate Dynamic Programming (ADP) algorithm that yields policies indicating which trials to be performed at each decision point and state.

2.4.3 Remarks

Optimising under uncertainty focuses on yielding a solution that performs well under a range of uncertainty realisations based on mathematical programming. Scenario-based approaches emerge to preserve model linearity and capture information on probability distribution however come with a limit in scalability. Once solved, the solutions of such approaches carry important value and help identify first stage decisions that perform well under uncertainty, corresponding to network designs or *here-and-now* initial decisions to be made.

Augmenting such frameworks with simulation tools may provide a range of answers, ranging from accuracy testing of optimisation solutions or an actual simulation of the dynamic performance of a network or policies adopted in manufacturing planning activities.

We find that the problem of capacity planning under demand uncertainty has been extensively studied and currently lacks the complexity related to these activities having to be performed in presence of process uncertainties as well. This is particularly relevant to planning ATMPs supply chains in early-stages of scale-up. By disregarding this initial uncertainty there is an increased risk of finding planning solutions which do not meet production targets, due to a worsened realised process performance, hence creating delays and shortages. Hence, this opens scope for identifying ways to plan under process and demand uncertainty simultaneously seeking to improve supply chain resilience to quantifiable process uncertainty.

With the aim of keeping the problem tractable, simulation based approaches have a great potential to support testing of accuracy of scenario-based optimisation as well as later on be coupled with optimisation to perhaps generate instances of *unknown* outcomes not considered by the optimisation. This would help quantify supply chain resilience in a more comprehensive fashion.

2.5 Multi-objective optimisation

Multi-objective optimisation allows the reconstruction of Pareto frontiers to understand conflicts and support decision-makers in the planning. In the context of cost, social, environmental sustainability, trade-offs may emerge depending on the studied systems. Some works in the literature have also considered the trade-off of these indicators with decisions risk under uncertainty. Below, we review works that have highlighted important decision-making trade-offs to consider when implementing the developed planning frameworks.

2.5.1 Emerging trade-offs

The first key trade-off emerges between cost and service level, emerging from a comparison of centralised versus decentralised systems. Liu and Papageorgiou (2013) compare costs, service levels and flow time of different capacity expansion

strategies. Overall, total costs decrease as sales are lost and increase as the total flow time increases (units distributed per unit time). However, cumulative capacity expansion helps mitigate total lost sales and thus minimise costs per unit sold. In Mestre et al. (2015), cost increases as access to medicines increases. By comparing planning strategies, the authors seek solutions that reduce costs per unit delivered. In Lakhdar et al. (2007), multi-objective optimisation is used to reconstruct conflicts between cost, service level and capacity utilisation and comparing cost different policies. Another interesting trade-off has been highlighted in Liu et al. (2014), where the anticipating maintenance planning reduces profits but prevents further down-time down the line. In Tsang et al. (2007b) the authors quantify the trade-off between downside risk and NPV and find that as the downside risk decreases the expected NPV also decreases, as the most profitable scenarios are potentially the most risk-taking ones too.

The integration of environmental sustainability may give rise to cost vs environment trade-offs to consider at the network design stage. Effectively, sustainable supply chain design and planning has been long studied in the PSE and operations research literature (Barbosa-Póvoa et al., 2018). Key contributions from Hugo and Pistikopoulos (2005) and Guillén-Gosálbez et al. (2010) is the introduction in the optimisation of LCA-based environmental constraints in the formulation which allow the reconstruction of trade-offs between costs and sustainability at the strategic level. This helps identify that they are conflicting in this context, embedding impacts from transport activities and storage activities. Elias Mota et al. (2020) highlight an important trade-off between the environmental impact and the NPV. However, screening different strategies helps identify the curves which achieve the same NPV with different costs. Duarte et al. (2022) use multi-objective optimisation to explore the trade-offs between cost, accessibility to vaccines and environmental footprint. They analyse supply chain structures obtained based on the order in which metrics are prioritised in the lexicographic method. For a given demand, increasing costs lead to increased environmental footprints, mainly resulting from transportation modes, which highlights the benefit for supply chain decentralisation. The same observation is made by Badejo and Ierapetritou (2024), who demonstrate that total supply chain costs increase with a minimisation of environmental footprints mainly from transport activities through decentralisation.

Environmental footprint indicators based on LCA are often embedding con-

flicting mid-point and endpoint impacts within. Sabio et al. (2012) address this multi-variate characteristic of sustainability using principal component analysis (PCA) to gain an understanding on the most impactful environmental categories. Overall, there again emerges a trade-off between cost and sustainability, driven primarily by transport.

2.5.2 Remarks

As one would expect, the larger the demands to satisfy through capacity planning, the larger the total costs. In pharmaceutical contexts, the risk of clinical trials failing and demands not realising challenges the investment planning, introducing a further trade-off between cost-social sustainability and the risk decision-makers are willing to take.

To add to this complexity, some works have highlighted a trade-off between total supply chain costs and environmental footprints when operating centralised or decentralised supply chains, as the latter decrease transport emissions whilst requiring more capacity investments. So far, no work has explored whether sustainability considerations in pharmaceutical capacity planning lead to alternative manufacturing solutions and create further dimensions to the trade-offs. Thus, it becomes worth exploring through multi-objective approaches whether cost-efficiency and environmental sustainability are in conflict or can work towards the same goal of improved manufacturing resource utilisation and product supply to meet target demands. To this end, superstructure-based optimisation problems can be formulated to explore manufacturing configurations, with a focus on investigating how operating modes (batch vs continuous), equipment types (SU vs MU) or product types impact both cost and sustainability.

2.6 Contribution of this work

Through this literature review we have identified that extensive work has been done in developing planning tools that consider decision-making at different scales and considering a range of uncertainties. Within the context of biopharmaceuticals, the problem of capacity planning has been thoroughly studied, with frameworks that embed complex scheduling models or simplify operations based on the application.

The first knowledge gap that emerges is tied to process uncertainty. To the best of our knowledge, there is no work that considers process uncertainties at the investment planning level. It is, therefore, worth investigating whether strategic planning improves if we consider a range of outcomes of process uncertainty. This is crucial to assess, given that most emerging sectors are undergoing a rapid scale-up before process capabilities are known for certain. Research questions that stem are tied to understanding the time-scales of this uncertainty and the meaning of it as process capabilities become gradually known over time and how can these be combined with demand uncertainty.

The second knowledge gap that emerges is that optimising under uncertainty is related to identifying solutions that perform well with respect to *known* unknowns which are uncertainties *a priori*. There is an opportunity to quantify measures of supply chain flexibility to withstand *unknown* unknowns which are unforeseen uncertainties.

The third knowledge gap relates to embedding sustainability in the decision-making, quantifying the magnitude of footprints of capacity planning on the environment and identifying the most sustainable alternatives. This should hold true for deterministic contexts and when decisions are made under uncertainty. Although the integration of sustainability has been considered for other industries and supply chain design problems, there is an opportunity to appended sustainability metrics in the context of pharmaceutical manufacturing scale-up.

In this thesis, these knowledge gaps are addressed in the following chapters:

1. Integration of process uncertainty: Chapter 3, 4, 5 and 6
2. Quantification of flexibility to unknown unknowns: Chapter 5
3. Integration of sustainability: Chapter 7

This thesis presents planning tools that embed detailed information on the economic, environmental and social sustainability of next-generation ATMPs manufacturing networks. Ultimately, the performance under uncertainty of any optimisation solution is the most valuable indicator of actual supply chain sustainability. Hence, we focus on developing *uncertainty-aware* decision-support tools to optimise networks to perform well upon realisation of both predictable and unforeseen uncertainties.

Chapter 3

Quantification of process uncertainty

Outputs of this chapter have been disseminated in:

- **Sarkis, M.**, Shah, N. & Papathanasiou, M.M. (2023). Characterization of key manufacturing uncertainties in next-generation therapeutics and vaccines across scales, *Journal of Advanced Manufacturing and Processing* 5(3), e10158.

3.1 Introduction

The growing demand for gene therapies and the need for vaccine manufacturing scale-up has highlighted bottlenecks in the global capacity for Advanced Therapy Medicinal Products (ATMPs). In the case of viral vectors-based gene therapies, shortages were reported as early as 2018 (Cytiva, 2021) and were further exacerbated by the recent pandemic as manufacturing capacity was repurposed towards urgent viral vector-based vaccine production. This highlighted that the viral vector industry is still immature when it comes to proactive and informed decision-making with respect to manufacturing capacity planning.

The increased demand and regulatory landscape pressures manufacturers to design cost-effective and reliable processes early on to guarantee global supply. In the context of an emerging manufacturing sector as that of viral vectors, this is not a straightforward exercise as product titers and recoveries tend to be lower compared to other biological products (Merten et al., 2014; Perry and Rayat, 2021; Moleirinho et al., 2019). The scale-up of cell-based biological processes is coupled with challenges related to the choice of operating conditions that ensure high product yields, cell growth and acceptable product quality. Currently, manufacturers are focusing on optimising their processes and are increasingly pressured to scale-up in presence of process uncertainties, which challenge production planning activities. To this end, computer aided tools enable a systematic analysis of correlations between underlying manufacturing uncertainties and key performance indicators (KPIs).

In Chapter 2, an extensive literature review of performance quantification frameworks has been presented, highlighting applications in biologics and ATMPs. In the latter studies, techno-economic models are complemented by uncertainty and sensitivity analyses to account for underlying uncertainties for processes under development.

Comisel et al. (2021b,a) developed a techno-economic model for lentivirus (LV) and adeno- associated virus (AAV) manufacturing and assessed how inputs related to material costs and resource costs impact cost of goods sold (COGs) per dose, using local sensitivity analysis (LSA). Ferreira et al. (2021) presented a bioprocess simulation of an adenoviral (AdV)-based vaccine production to asses alternative upstream process (USP) configurations, including batch versus perfusion mode of operation. The study proposed a what-if scenario approach to assess

how USP titer impacts capital and operating expenditures comparing nominal process performance, at different scales. Focusing on COVID-19 vaccines, Kis et al. (2022) assessed the impact of manufacturing uncertainties during process development through global sensitivity analysis (GSA), including batch failure rate, process scale, production titers, resource costs and their impact on process capital expenditure (CapEx), operating expenditure (OpEx) and throughput-related quantities. In this case, production scale is integrated as a continuous input, quantifying total production capacities and investments needed to deliver billions of COVID-19 vaccines and booster doses at global scale.

Despite previous contributions in the field, the scale-dependent quantification of manufacturing uncertainty remains an open challenge and is crucial to develop informed scale-up strategies. To address this, in this Chapter we present a model-based framework to quantify and characterise process performance uncertainties. Specifically, we assess how propagation of underlying manufacturing uncertainties affects investment and production planning performance indicators, such as batch size, process bottlenecks and costs. Thus, we capture the impact of volumetric scale up and compare worst-case and best-case scenario performances at different scales. The former case is accounted for by sampling for lower viral titers, lower recoveries, and longer tasks, while the latter is quantified by sampling for improved titers, higher recoveries, and shorter process times.

The proposed framework is applied to three industrially relevant viral vectors, currently at the forefront of ATMPs manufacturing for vaccine and gene therapy applications. The framework is transferable to any product of interest and can be used by technology developers and manufacturers to assess process feasibility and quantify uncertainty early on. Additionally, the quantification of product-specific and process-specific uncertainty is an important first step towards the development of optimisation-based planning strategies which integrate this acquired uncertainty knowledge and thus made *uncertainty-aware*.

3.2 Methodology

The framework presented in this Chapter includes three main steps: (i) data collection, (ii) flow-sheet modelling and simulation, (iii) uncertainty and sensitivity analyses. The methodologies underlying each of these steps are discussed in the following sections.

3.2.1 Data collection

Information regarding production processes for each of the viral vector manufacturing process is obtained from the scientific literature, for both drug substance (primary manufacturing) and drug product (secondary manufacturing). A schematic representation of the studied processes is presented in Figure 3.1, alongside a summary of the operations taking place in each of the process steps for AdV (Table 3.1), LV (Table 3.3) and AAV (Table 3.5).

Data collection also involves recording input uncertainty ranges, used for sensitivity and uncertainty analysis. Nominal values and ranges for process parameters and costs are listed in Tables 3.2 (AdV), 3.4 (LV) and 3.6 (AAV). Specifically, in this Chapter we consider a $\pm 50\%$ variation from nominal values, which correspond to values for each process parameters as reported in the literature and industrial reports. For some inputs, including cell culture tasks duration, titers and purification yields, more detailed information on parameter ranges is obtained from the literature and hence integrated as uncertainty range. Ultimately, the user can incorporate any knowledge on input-specific uncertainty distributions, if this is available.

3.2.2 Flowsheet development

Following data collection, process flowsheets are developed using SuperPro Designer (Intelligen). SuperPro uses built-in sets of algebraic and differential equations to calculate material and energy balances, sizes equipment, schedules operations, and computes capital and operating expenditures. SuperPro is suitable for modelling batch and continuous bioprocesses and incorporates an extensive industrial-based database for biopharmaceutical applications. The platform allows for kinetic or stoichiometric modelling of reaction steps. In this Chapter, we adopt the latter approach, as the literature reports extensive information on product titer and reaction durations. Furthermore, separation processes can be modelled in detail or via short-cut methods, based on the information available. Hence, we rely on shortcut-methods for chromatography and filtration unit models and collect data of step yields and step duration.

Based on the throughput-related outputs (Figure 3.2), economic calculations are performed. We consider processes for three viral vector platforms, including AdV, LV and AAV. For the purposes of these models, we assume processes that

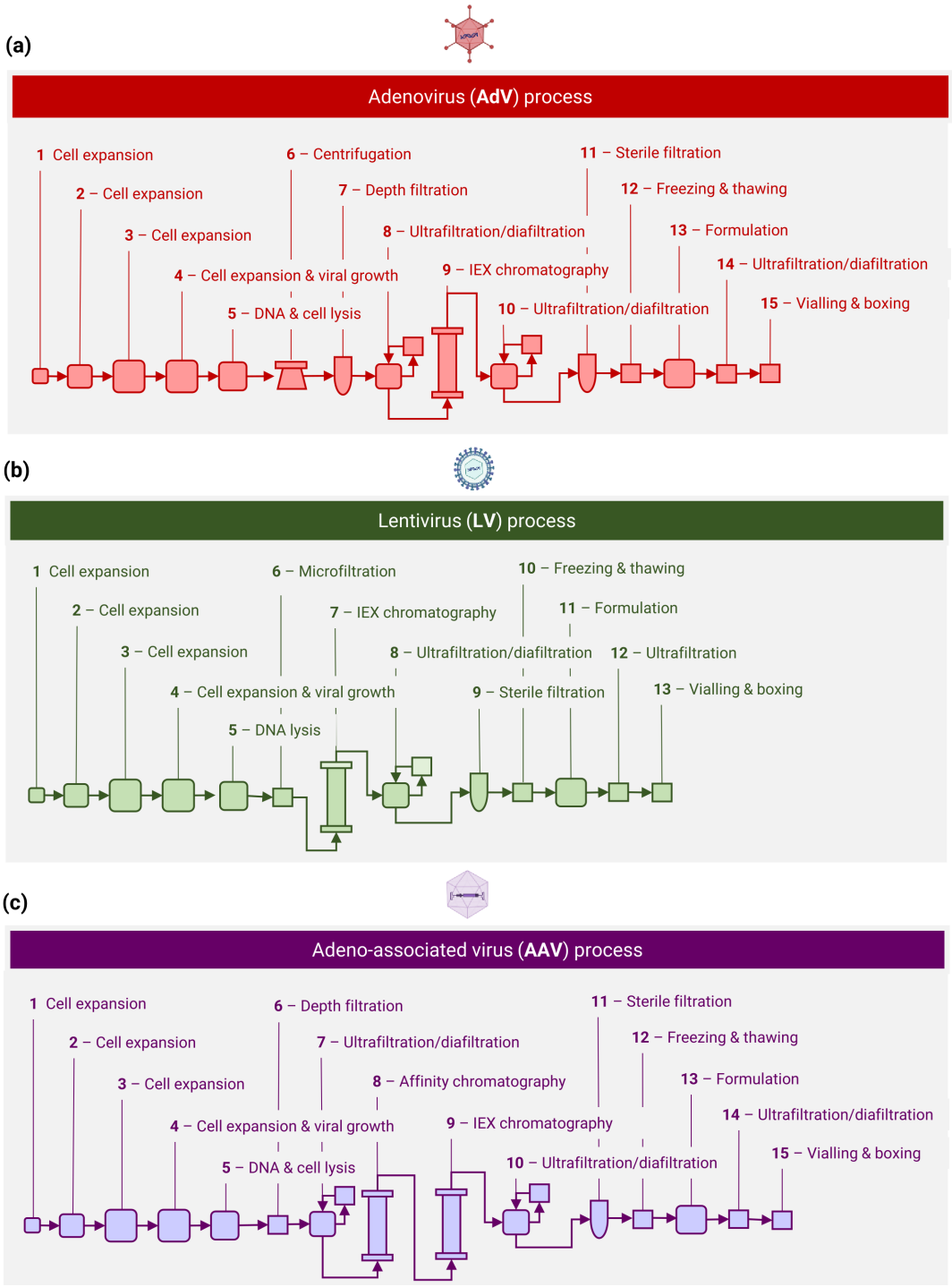


Figure 3.1: Flowsheet for: (a) Adv manufacturing highlighting labeled process steps (1–15) and key resource inputs, (b) LV manufacturing highlighting labeled process steps (1–13) and key resource inputs, and (c) AAV manufacturing highlighting labeled process steps (1–15) and key resource inputs.

Table 3.1: AdV process recipe. Process information was retrieved from Ferreira et al. (2021), Intelligen, Kis et al. (2022), Merten et al. (2014)

Step	**Operation**	**Start (h)**	**Duration (h)**	**Yield (VP/mL)**
1	Add fresh serum-free medium in 1:6 v/v dilution ratio	1.76	0	-
1	Seed HEK 293 cells to initiate cell growth	2.27	0	-
1	Cell expansion (batch) to target density 1.4×10^6 cells/mL	2.52	86.5	-
2	Add fresh serum-free medium in 1:6 v/v dilution ratio	98.02	0.5	-
2	Cell expansion (batch) to target density 1.4×10^6 cells/mL	99.02	86.5	-
3	Add fresh serum-free medium in 1:6 v/v dilution ratio	194.02	0.5	-
3	Cell expansion (batch) to target density 1.4×10^6 cells/mL	196.02	86.5	-
4	Add fresh serum-free medium in 1:6 v/v dilution ratio	288.92	2.85	-
4	Cell expansion (batch) to target density 1.9×10^6 cells/mL	292.73	86.5	-
4	Infect culture with AdV inoculum	388.73	0	-
4	Viral growth to target titer and 2.5×10^6 cells/mL	388.98	48	1×10^{11}
5	Add 1 mL of lysis buffer per 10^7 cells	440.54	0.89	1×10^{11}
5	Cell lysis (100%) to break down HEK293 cells	441.67	1	1×10^{11}
5	Add 25 U of Benzonase per 10^6 mL	442.67	0	1×10^{11}
5	DNA lysis	442.93	1	1×10^{11}
6	Centrifuge mixture to achieve target 300 g/L solids	443.93	4	0.9×10^{11}
7	Depth filter mixture at 720 L/m^{2}h	447.93	0.25	0.81×10^{11}
8	Concentrate by factor of 20 & diafilter	448.36	4	0.73×10^{11}
9	Load IEX column of capacity 3×10^{12} VP/mL	452.36	0.08	-
9	Wash IEX column with 10 BV	452.45	0.33	-
9	Elute product with NaCl solution gradient 0-5 g/L	452.78	0.07	1.2×10^{12}
9	Regenerate IEX column with 10 BV of high NaCl buffer	452.85	0.33	-
9	Equilibrate IEX column with 10 BV of buffer	453.18	0.33	-
10	Concentrate by factor 3	453.20	4	-
10	Diafilter & buffer exchange in 1:7 v/v ratio	453.20	4	2.1×10^{12}
11	Sterile filtration at 2000 L/m^{2}h	457.2	1	2.0×10^{12}
12	Freeze at -70°C & Store drug substance	457.9	0.46	-
13	Thaw drug substance at 25°C	458.1	0.45	-
14	Formulate drug substance with PBS buffer	458.8	0.09	-
14	Ultrafiltration	459	4	1×10^{11}
15	Fill 400 vials/min, cap and label vials	459	4	-
15	Refrigerate at 2°C and box	459	4	-

Table 3.2: Operational inputs for AdV flowsheet. Ranges are marked as known (Y), if supported by literature, or unknown (N), if ±50% of its nominal value is assumed. Rejection coefficients inputs to obtain the target recoveries are reported in parenthesis. Values retrieved from (a) Ferreira et al. (2021), (b) Jagschies et al. (2018), (c) Fedosyuk et al. (2019), (d) Farnós et al. (2020), (e) Cytiva

.

Input Name	**Step**	**Nominal**	**Known?**	**LB**	**UB**	**Units**	**Reference**
Ferment 1 task	1	4	Y	3	6	days	(a-b)
Ferment 2 task	2	4	Y	3	6	days	(a-b)
Ferment 3 task	3	4	Y	3	6	days	(a-b)
Ferment 4 task	4	4	Y	3	6	days	(a-b)
Viral growth task	4	2	Y	1	2	days	(a-c)
Viral titer	4	1×10^{11}	Y	5×10^{10}	3×10^{11}	VP/mL	(c-d)
Working/vessel volume	4	71	Y	20	80	%	(e)
Lysis task	5	1	N	0.5	1.5	h	(a)
Degrade DNA task	5	1	N	0.5	1.5	h	(a)
Centrifugation recovery	6	10	N	5	15	%	(a)
Centrifugation task	6	4	N	2	6	h	(a)
Filtration 1 recovery	7	10	N	5	15	%	(a)
Filtration 1 task	7	1	N	0.5	1.5	h	(a)
UF/DF1 recovery	8	90 (0.995)	Y	90 (0.995)	95 (0.996)	%	(a-c)
UF/DF1 task	8	4	N	2	6	h	(a)
IEX elution recovery	9	80	N	48	80	%	(a-c)
UF/DF2 recovery	10	90	Y	90 (0.987)	95 (0.995)	%	(a-c)
UF/DF2 task	10	4	N	2	6	h	(a)
Filtration 2 recovery	11	2	N	1	4	%	(a)
Filtration 2 task	11	1	N	0.5	15	h	(a)

Table 3.3: LV process recipe. Process information was retrieved from Ferreira et al. (2021), Intelligen, Perry and Rayat (2021), Comisel et al. (2021b)

Step	**Operation**	**Start (h)**	**Duration (h)**	**Titer (TU/mL)**
1	Add fresh serum-free medium in 1:6 v/v dilution ratio	1.76	0	-
1	Seed HEK 293 cells to initiate cell growth	2.27	0	-
1	Cell expansion (Batch) to target density 1.4×10^6 cells/mL	2.52	86.5	-
2	Add fresh serum-free medium in 1:6 v/v dilution ratio	98.02	0.5	-
2	Cell expansion (Batch) to target density 1.4×10^6 cells/mL	99.02	86.5	-
3	Add fresh serum-free medium in 1:6 v/v dilution ratio	194.02	0.5	-
3	Cell expansion (Batch) to target density 1.4×10^6 cells/mL	196.02	86.5	-
4	Add fresh serum-free medium in 1:6 v/v dilution ratio	288.92	2.85	-
4	Cell expansion (Batch) to target density 1.9×10^6 cells/mL	292.73	86.5	-
4	Transfect culture with plasmid (2.5 μg/10^6 cells)	388.73	0	-
4	Viral growth to target titer and 2.5×10^6 cells/mL	388.98	48	1×10^7
5	Add 25 U of Benzonase per 10^6 mL	443.54	0	-
5	DNA lysis (Conversion: 90%)	443.79	1	-
6	Microfiltration to remove cells and precipated nucleic acids	444.79	4	-
7	Load IEX column wit binding capacity: 5×10^8 TU/mL	445.79	7.62	-
7	Wash IEX column with 10 BV of buffer	453.4	0.33	-
7	Elute product with NaCl solution gradient 0-5g/L	453.74	0.14	1.7×10^8
7	Regenerate IEX column with 10 BV of high NaCl buffer	453.88	0.33	-
7	Equilibrate IEX column wit 10 BV of buffer	454.22	0.33	-
8	Concentrate by factor of 6	454.0	4	-
8	Diafilter & buffer exchange in 1:7 v/v ratio	454.0	4	9.8×10^8
9	Sterile filtration at 2000 L/m^{2}h	458.0	1	8×10^8
10	Freeze & store drug substance at -70°C	459.0	0.46	-
10	Thaw drug substance to 25°C	459.4	0.45	-
11	Formulate drug substance with PBS buffer	459.9	0.09	-
12	Ultrafiltration	461.0	4	2×10^9
13	Fill 10 vials/min (1-5 doses/vial), cap and label vials	461.0	4	-
13	Cryofreeze and box at -70°C	461	4	-

Table 3.4: Operational inputs for LV flowsheet. Ranges are marked as known (Y), if supported by literature, or unknown (N), if ±50% of its nominal value is assumed. Rejection coefficients inputs to obtain the target recoveries are reported in parenthesis. Values retrieved from (a) Ferreira et al. (2021), (b) Jagschies et al. (2018), (c) Fedosyuk et al. (2019), (d) Comisel et al. (2021b), (e) Ansorge et al. (2009), (f) Bauler et al. (2020b), (g) Merten et al. (2011), (h) Merten et al. (2014), (i) Cytiva, (j) Perry and Rayat (2021), (k) Intelligen.

Input Name	**Step**	**Nominal**	**Known?**	**LB**	**UB**	**Units**	**Reference**
Ferment 1 task	1	4	Y	3	6	days	(a-b)
Ferment 2 task	2	4	Y	3	6	days	(a-b)
Ferment 3 task	3	4	Y	3	6	days	(a-b)
Ferment 4 task	4	4	Y	3	6	days	(a-b)
Viral growth task	4	2	Y	1	2	days	(a-c)
Viral titer	4	1×10^7	Y	2×10^6	1.5×10^8	TU/mL	(d-g)
Working/vessel volume	4	71	Y	20	80	%	(i)
Microfiltration recovery	5	80 (0.45)	N	70 (0.6)	90 (0.2)	%	(d)-(j)
Microfiltration task	5	4	N	2	6	h	(k)
Degrade DNA task	6	1	N	0.5	1.5	h	(a)
IEX elution recovery	7	40	Y	30	90	%	(d)-(j)
UF/DF recovery	8	80 (0.974)	Y	70 (0.96)	90 (0.987)	%	(d)-(j)
UF/DF task	8	4	Y	2	6	h	(k)
Filtration recovery	9	25	N	13	38	%	(a)
Filtration task	9	1	N	0.5	15	h	(k)

Table 3.5: AAV process recipe. Process information was retrieved from Ferreira et al. (2021), Intelligen, Cameau et al. (2019), Comisel et al. (2021a), Qu et al. (2015)

Step	**Operation**	**Start (h)**	**Duration (h)**	**Titer (vg/mL)**
1	Add fresh serum-free medium in 1:6 v/v dilution ratio	1.76	0	-
1	Seed HEK 293 cells to initiate cell growth	2.27	0	-
1	Cell expansion (batch) to final density of 1.4×10^6 cells/mL	2.52	86.5	-
2	Add fresh serum-free medium in 1:6 v/v dilution ratio	98.02	0.5	-
2	Cell expansion (batch) to final density of 1.4×10^6 cells/mL	99.02	86.5	-
3	Add fresh serum-free medium in 1:6 v/v dilution ratio	194.02	0.5	-
3	Cell expansion (batch) to final density of 1.4×10^6 cells/mL	196.02	86.5	-
4	Add fresh serum-free medium in 1:6 v/v dilution ratio	288.92	2.85	-
4	Cell expansion (batch) to final density of 1.9×10^6 cells/mL	292.73	86.5	-
4	Transfect culture with 2.5 g/10^6 cells	388.73	0	-
4	Viral growth to target titer and final density 2.5×10^6 cells/mL	388.98	48	1×10^{11}
5	Add 1 mL lysis buffer per 10^7 cells	437.4	0.8	1×10^{11}
5	Cell lysis	438.4	1.5	1×10^{11}
5	Add 25 U of Benzonase per 10^6 mL)	439.8	0	1×10^{11}
5	DNA lysis	440	1	1×10^{11}
6	Microfilter and concentrate by 20	441	4	1.2×10^{11}
7	Concentrate by factor of 20	446	4	1.2×10^{11}
7	Load affinity column with binding capacity of 3×10^{12} vg/mL	450	0.2	-
7	Wash affinity column with 10 BV of buffer	450	0.2	-
7	Elute product with NaCl solution gradient 0-5 g/L	450	0.1	1.2×10^{12}
7	Regenerate affinity column with 10 BV of buffer	450.2	0.3	-
7	Equilibrate affinity column with 10 BV of buffer	450.5	0.3	-
8	Load IEX column with binding capacity 1×10^{13} vg/mL	450.5	0.5	-
8	Wash IEX column with 10 BV of buffer	450.5	0.3	-
8	Elute product with NaCl solution gradient 0-5 g/L	450.8	0.07	1.12×10^{12}
8	Regenerate IEX column with 10 BV of high NaCl buffe	450.9	0.3	-
8	Equilibrate IEX column with 10 BV of buffer	451.2	0.3	-
9	Concentrate by factor of 3	450.5	5	1.12×10^{12}
10	Diafilter	450.5	5	-
11	Sterile filtration at 2000 L/m^2h	451.0	1	2.0×10^{12}
12	Freeze and store drug substance at -70°C	452.0	0.46	-
13	Thaw drug substance to 25°C	452.4	0.45	-
13	Formulate drug substance with PBS buffer	452.9	0.09	1×10^{13}
14	Ultrafiltration	454.0	4	-
15	Fill 40 vials/min, cap, and label vials	454.0	4	-
15	Cryofreeze and box at -70°C	454.0	4	-

Table 3.6: Operational inputs for AAV flowsheet. Ranges are marked as known (Y), if supported by literature, or unknown (N), if ±50% of its nominal value is assumed. Rejection coefficients inputs to obtain the target recoveries are reported in parenthesis. Values retrieved from (a) Ferreira et al. (2021), (b) Jagschies et al. (2018), (c) Qu et al. (2015), (d) Thorne et al. (2009), (e) Cecchini et al. (2011), (f) Grieger et al. (2016), (g) Comisel et al. (2021a), (h) Cytiva

Input name	**Step**	**Nominal**	**Known?**	**LB**	**UB**	**Units**	**Reference**
Ferment 1 task	1	4	Y	3	6	days	(a-b)
Ferment 2 task	2	4	Y	3	6	days	(a-b)
Ferment 3 task	3	4	Y	3	6	days	(a-b)
Ferment 4 task	4	4	Y	3	6	days	(a-b)
Viral growth task	4	2	Y	1	2	days	(a-c)
Viral titer	4	1×10^{11}	Y	5×10^{10}	1×10^{11}	vg/mL	(c-g)
Working/vessel volume	4	71	Y	20	80	%	(h)
Cell lysis task	5	1	N	0.5	1.5	h	(a)
Degrade DNA task	5	1	N	0.5	1.5	h	(a)
Microfiltration recovery	6	95 (0.05)	Y	90 (0.1)	95 (0.05)	%	(c)-(g)
Microfiltration task	6	4	N	2	6	h	(a)
UF/DF1 recovery	7	95 (0.995)	Y	90 (0.99)	95 (0.995)	%	(a)-(c)-(g)
UF/DF1 task	7	4	N	2	6	h	(a)
Affinity elution recovery	8	70	Y	70	94	%	(c)-(h)
IEX elution recovery	9	70	Y	70	94	%	(c)-(h)
UF/DF2 recovery	10	95 (0.995)	Y	90 (0.993)	95 (0.995)	%	(a)-(c)-(g)
UF/DF2 task	10	4	N	2	6	h	(c)-(h)
Recovery SF	11	25	N	13	38	%	(a)
Filtration task	11	1	N	0.5	15	h	(a)

Table 3.7: Resource costs inputs for AdV, LV and AAV flowsheets. Ranges of inputs are marked as known (Y) or unknown (N). The former implies that the range is supported by literature, the latter implies that range was calculated as ±50% of its nominal value. Values retrieved from (a) Ferreira et al. (2021), (b) Comisel et al. (2021b), (c) Comisel et al. (2021a), (d) Intelligen, (e) Kis et al. (2022)

Input Name	**Process**	**Nominal**	**Known?**	**LB**	**UB**	**Units**	**Reference**
Medium cost	AdV, LV, AAV	40	N	30	60	$/L	(a)
pDNA cost	LV, AAV	74,000	N	37,000	110,000	$/g	(b-c)
Endonuclease cost	AdV, LV, AAV	12,000	N	6,000	24,000	$/5MU	(b-c)
Lysis buffer cost	AdV, AAV	55	N	27.5	82.5	$/kg	(a)
Chrom. buffer cost	AdV, LV, AAV	0.1	N	0.05	0.15	$/L	(d)
Resins cost	AdV, LV, AAV	1,500	N	750	2,250	$/L	(d)
Vials cost	AdV, LV, AAV	2	N	1	3	$/vial	(d-e)
MF membrane cost	AdV, LV, AAV	400	N	200	600	$/m^2	(d)
UF membrane cost	AdV, LV, AAV	981	N	490.5	1,471.5	$/m^2	(d)
SF membrane cost	AdV, LV, AAV	1,000	N	500	1,500	$/item	(d)
labour basic rate	AdV, LV, AAV	30	N	15	45	$/h	(d)
QC % (of labour cost)	AdV, LV, AAV	30	N	15	45	%	(d)

have already been optimised and approved by regulatory authorities. Hence, key process parameters, such as temperatures, pressures, medium, and buffer compositions are fixed.

As can be seen in Tables 3.1-3.5, major differences in the processes are related to viral growth and the downstream of the processes (DSP). The first key difference stems from the fact that AdV propagation is usually initiated with infection (Merten et al., 2014) rather than plasmid transfection. The purification strategy differs due to AdV and AAV being intracellular products thus requiring cell lysis after harvesting for the viral particles to be released in the supernatant (Merten et al., 2014), which is not required for LV particles (Perry and Rayat, 2021). Further details on the modelling is presented in Appendix A.

3.2.2.1 Outputs

Key performance indicators (KPIs) can be divided in throughput- and cost- related outputs (Figure 3.2a). The batch size χ of each process is calculated based on the mass flow per batch. The assumed dose sizes relate to the 3 selected representative product applications, namely vaccines as 5×10^{10} AdV viral particles (VP)/dose CHMP (2021a), in vivo gene therapy as 7×10^{14} AAV viral genomes (vg)/dose CHMP (2018), ex vivo gene therapy as 2×10^{9} LV transfecting units

(TU)/dose Comisel et al. (2021b). The batch time α is calculated as the time to process the batch from USP to fill and finish (F&F). The cycle time τ is defined as the minimum required time before another batch can be started (i.e. bottleneck step).

The capital cost includes direct fixed capital (DFC) costs C_1, start-up and validation C_2 and working capital costs C_3. DFC includes equipment purchase cost and estimates costs of installation, piping, and building. Start-up and validation cost is calculated as a percentage of DFC Intelligen. Working capital is instead defined as the upfront expenditure required to cover the first 30 days of operation in terms of variable costs. It therefore depends on the variable-related OpEx per batch (materials O_1, consumables O_4, utilities O_5) and how many batches can be completed within the starting 30 days.

The operating cost includes raw material costs (medium, buffers, lysing agents, plasmid DNA, and HEK cells), consumables (chromatography resins, filter membranes), utilities, labour costs, QC/QA costs and facility dependent costs. The variable costs (materials O_1, consumables O_4, utilities O_5) are calculated based on resource costs reported in the literature and SuperPro industrial-based materials database Table 3.7. The labour costs O_2 are calculated based on an assumed basic labour rate 30 \$/h (Intelligen) and on the calculated batch time τ.

The assignment of labour to each process step can be specified by an allocation factor, which relates the number of labour hours to the actual process time. A value ≥ 1 would indicate that operators would work in parallel on a specified task. In the presented case studies, a value of 1 is assumed for all units. The quantified labour hours are multiplied by a lumped labour rate, which is calculated based on the input basic rate and accounting for benefits, supervision, supplies and administration requirements. QC/QA costs are calculated as a percentage of this. Finally, facility dependent costs O_3 are calculated as a percentage of DFC and account for the contribution of maintenance and equipment depreciation to the operating costs. In this work, OpEx per batch and OpEx per dose are selected as outputs.

Different scales of production are considered for each vector product, namely 50, 200, 1000, and 2000 L scale (main bioreactor working volume). The throughput of the facility is adjusted via a user-defined factor which scales flows in each material balance calculation. Process models are simulated for the selected scales and products, thereby solving material balance, equipment sizing and cost-

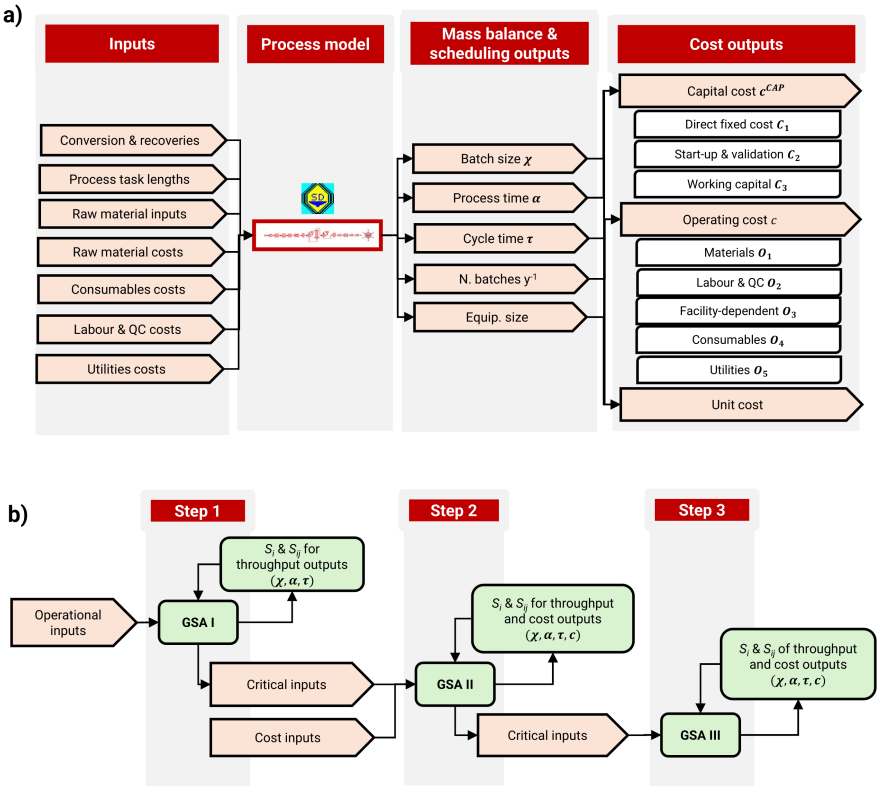


Figure 3.2: (a) Overview of layers of calculations embedded in techno-economic model and input-output space overview for GSA and UA and (b) Sequential GSA workflow. Step 1: highest-ranking ($S_i + S_{ij} \geq 0.05$) operational inputs are obtained, which impact throughput-related outputs. Step 2: the screen operational inputs are combined with resource costs inputs. GSA determines the highest-ranking inputs within the operational and resource costs input space. Outputs monitored in GSA II include capital costs and operating costs. Finally, GSA III is performed.

ing equations. The techno-economic performance of each process plant at this nominal value is recorded and the cost breakdown is analysed. The outputs include overall process throughput, time, capital, and operating costs. The output set can be augmented to section-specific quantities and/or individual cost components. Section-specific performance quantification directly ties in optimisation formulations (Chapter 4, 5 and 6) considering multi-product manufacturing configurations.

3.2.3 Uncertainty analysis

Uncertainty analysis (UA) is performed to quantify output uncertainty distributions and allow comparisons amongst processes. To align with current industrial

practices, where campaign planning and process improvements take place post the establishment of the process units, equipment size is fixed for all process models.

The framework uses a SobolGSA-MatLab-Component Object Model (COM)-SuperPro interface, 1000 quasi-random Sobol-based input samples and iteratively simulates and records outputs for each sample. Processes are assumed to be *independent* and UA is conducted on each process flowsheet separately. In addition to this, the inputs of each flowsheet are assumed to be independent and Sobol' sampling is used to achieve high coverage of the input uncertainty space.

By considering a range of operational input realisations whilst process design and scale is fixed, the framework effectively helps quantify a *productivity space* for each of the studied processes. This productivity may decrease across scales due to limited understanding of the volumetric scale up or improve with process engineering and optimisation. Our sampling approach captures this and helps a comparison amongst the independent processes, disregarding potential correlations. The variability of the KPIs is assessed by computing median and interquartile ranges (IQR). The latter indicates the range covered by the central 50% data points within the collected output samples. The impact of assumed input distributions on the output ranges is also studied.

To cater for all steps across process development and commercial scale, we propose the investigation of two types of uncertainties. The first one can be understood as the *epistemic* uncertainty, considering significant lack of process understanding and therefore uniform distributions for the inputs. The second type of uncertainty would correspond the *aleatoric* uncertainty, whereby process knowledge has been established and therefore the most likely input value can be identified. To assess such cases, we use triangular distributions based on reported upper and lower input bounds and a most probable value which corresponds to the nominal input considered for the techno-economic assessment.

3.2.4 Global sensitivity analysis

The UA provides a relation between input-output uncertainty spaces. The contribution of input variability on output variables can be quantified by via global sensitivity analysis (GSA) for each process considered. This step requires a larger number of samples for sensitivity calculations to be performed. The techno-

economic models are simulated iteratively for 4096 quasi-random input samples leveraging on the development of SobolGSA-MatLab-Component Object Model (COM)-SuperPro interface. This allows an investigation of the respective contributions of input uncertainty to the recorded outputs. Given the dimensionality of the input–output model, random sampling-high dimensional model representation (RS-HDMR) is implemented. The method relies on the generation of a metamodel from input–output data and requires the additional samples (1023) to simulate the process and the test the validity of the surrogate.

By varying all inputs simultaneously GSA computes indices which inform on both single input contributions to output variability (first-order indices, S_i), combined contributions of inputs i and j (second-order indices, S_{ij}), as well as higher order interactions. The total contribution of a single input (S_T) is also calculated, by summing up first, second and higher orders interactions. Here we assume first-, and second-order interactions and calculate: $S_T = S_i + S_{ij}$. Higher order interactions are excluded as insignificant.

Given the computational complexity of considering all operational inputs and resource inputs, a sequential GSA-based methodology is proposed to screen inputs progressively and identify critical sensitivities (Figure 3.2b). In doing so, we take advantage of the layers of calculations, namely material balances, scheduling and economic calculations (Figure 3.2a). First, operational inputs for each manufacturing process are ranked (GSA I), considering only batch size χ, process time α, and cycle time τ as outputs. Inputs i where the condition $S_T \leq 0.05$ are deemed to have a negligible contribution to output variability and excluded from the input sets. The first reduced set of operational inputs is combined with the full set of resource cost inputs (GSA II) to determine the most impactful inputs on cost-related quantities. Additionally capital costs c^{CAP}, operating costs c^{OP} and costs per dose are considered and a second reduced set of inputs is obtained based on the selected threshold of significance and simulated via GSA III. The knowledge acquired via this step provides an understanding of input–output sensitivities and assists the interpretation of the UA.

For the purposes of this Chapter, we perform GSA to compare the three processes at the same scale (2000L). However, this can be repeated for any process of interest and assumed input distributions.

3.3 Results & discussion

3.3.1 Techno-economic analysis of vector processes

Each of the developed flowsheets is simulated in SuperPro with all inputs fixed at the base case values to determine nominal throughputs and costs of each manufacturing platform for process comparison. Capital costs attributed to seed cell expansion steps (1–3) (Figure 3.1a–c) are estimated to the be same for the three processes at 40M$ (Figure 3.3a-c). Although equipment dimensions and direct fixed cost correspond to 62M$ for all processes in the virus production section as well, LV and AAV require an additional upfront working capital 1.6M$ related to plasmid-based transfection. Upfront working capital accounts for 30days of operation (OpEx per month) and the raw materials in this section (plasmid DNA) make up approximately 77% of the OpEx per batch contribution for the virus propagation process section (Figure 3.3d and f).

The primary recovery and clarification steps require similar equipment purchase costs for AdV and AAV 11–12M$, which result higher than equipment costs LV (3M$), as the supernatant can be harvested, microfiltered and degraded without UF steps. LV are extracellular entities and there is not an additional expenditure in lysis buffers, hence reducing the OpEx per batch brought by raw materials. In the case of AdV and AAV instead, lysis and Benzonase share $1.9 \times 10^5$$/batch of the OpEx per batch (Figure 3.3b and f).

The purification train for AdV and LV involves ion-exchange chromatography (IEX) and ultra-filtration/diafiltration (UF/DF) steps. AAV purification section involves an additional affinity chromatography step which comes with additional consumables costs and facility-dependent costs due to maintenance and equipment depreciation (Figure 3.3a and f). In this step, a smaller bed volume is required compared to the previous IEX step as volumetric flowrates handled are lower. Capital costs in the polishing step result similar across all processes (8.6–8.8 M$), as overall small deviations in the stream volumetric flowrates across processes.

The F&F steps of the process entail different equipment sizes, because of the product-specific dilution targets in each vial. With a vial price at 2$/vial, the AdV product is more diluted and is filled in smaller vials (5mL). This results in a need for equipment at higher throughput (maximum 400 vials/min), in three

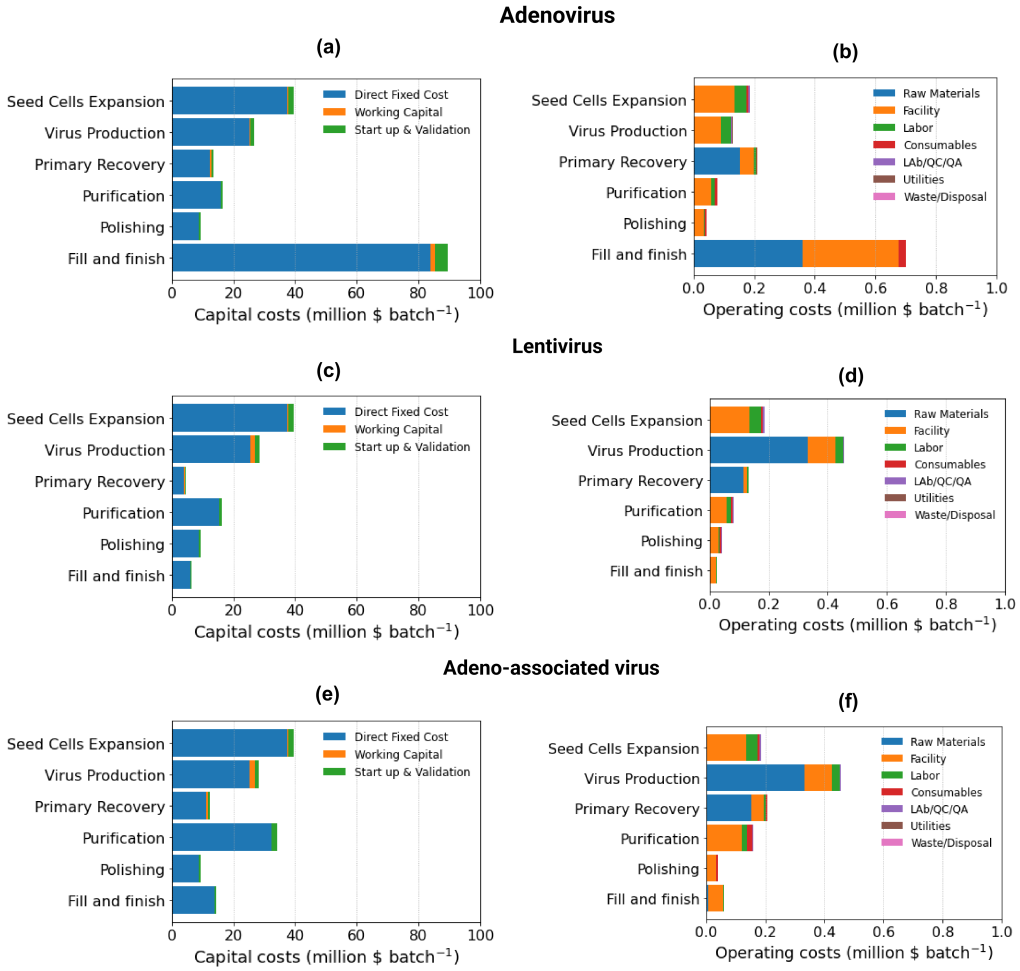


Figure 3.3: Breakdown of costs of manufacturing for viral vector processes: (a) capital costs for AdV process, (b) operating costs for AdV process, (c) capital costs for LV process, (d) operating costs for LV process, (e) capital costs for AAV process, (f) operating costs for AAV process.

parallel lines and many vials required. This translates into a substantial raw material costs (3.3×10^4\$/batch), as can be seen in Figure 3.3A. An additional cost contributor is the maintenance and depreciation cost, which directly depends on the capital cost. AAV and LV gene therapies require equipment of capacity $<$ 25 units/min, bypassing the need for parallel lines and requiring a smaller number of vials as products are more concentrated.

Production times for each batch range around 22 days, with 95% of this contributed by the USP process and DSP time taking 20.5 days for each of the processes, respectively. As labour cost is estimated as a function of a basic labour rate and the hours of operation per batch, the labour cost contribution

to OpEx per batch results similar for all processes. A process bottleneck of 6 days is recorded for all processes which corresponds to the cell culture step and viral growth in the main bioreactor and corresponds to 6 days. The maximum throughput of each facility results to be the same, namely 50 batches/year assuming a single year-long production campaign.

Process batch sizes that result from each process are of significantly different orders of magnitude. The AdV process produces a nominal value of 3.0×10^6 doses/batch of AdV, compared to 214 LV doses/batch and 102 AAV doses/batch. The resulting nominal costs per dose are 0.5 $/dose for the AdV vaccine, 11600 $/dose of the AAV *in vivo* gene and 1650 $/dose for LV products for *ex vivo* gene therapy. The results are found to agree with the orders of magnitudes for batch sizes and costs per dose reported in literature (Comisel et al., 2021b,a; Kis et al., 2022; Cameau et al., 2019; Peter Jones, 2019). Specifically, this indicates that vaccine products are consistently cheaper than target therapeutics, which is to be expected given the product-specific dose sizes and production yields.

Whilst some process steps are product-specific, overall viral vector production platforms are similar (e.g. USP train, IEX chromatography, filtration). Equipment-related differences, which propagate to facility-related costs, relate to the F&F step for AdV manufacturing, which processes a significantly more diluted product, and the additional affinity chromatography step in AAV manufacturing, for further polishing. In addition to this, key cost differences stem from different raw materials being used for the same process steps, In practice, this means that there is an opportunity to design multi-product facilities with some product-specific equipment and raw materials. Ultimately, this flexibility would also come with an inherent trade-off between using established capacity for large-scale production of low cost vaccines and costly medicines which target a subpopulation.

3.3.2 Uncertainty quantification

3.3.2.1 Throughput outputs

Figure 3.4 gathers the results of the UA considering a range of process scales and different input distributions, including uniform and triangular distributions.

A volumetric propagation of uncertainty is observed with respect to the batch size for all production platforms and distributions, with larger IQR ranges as

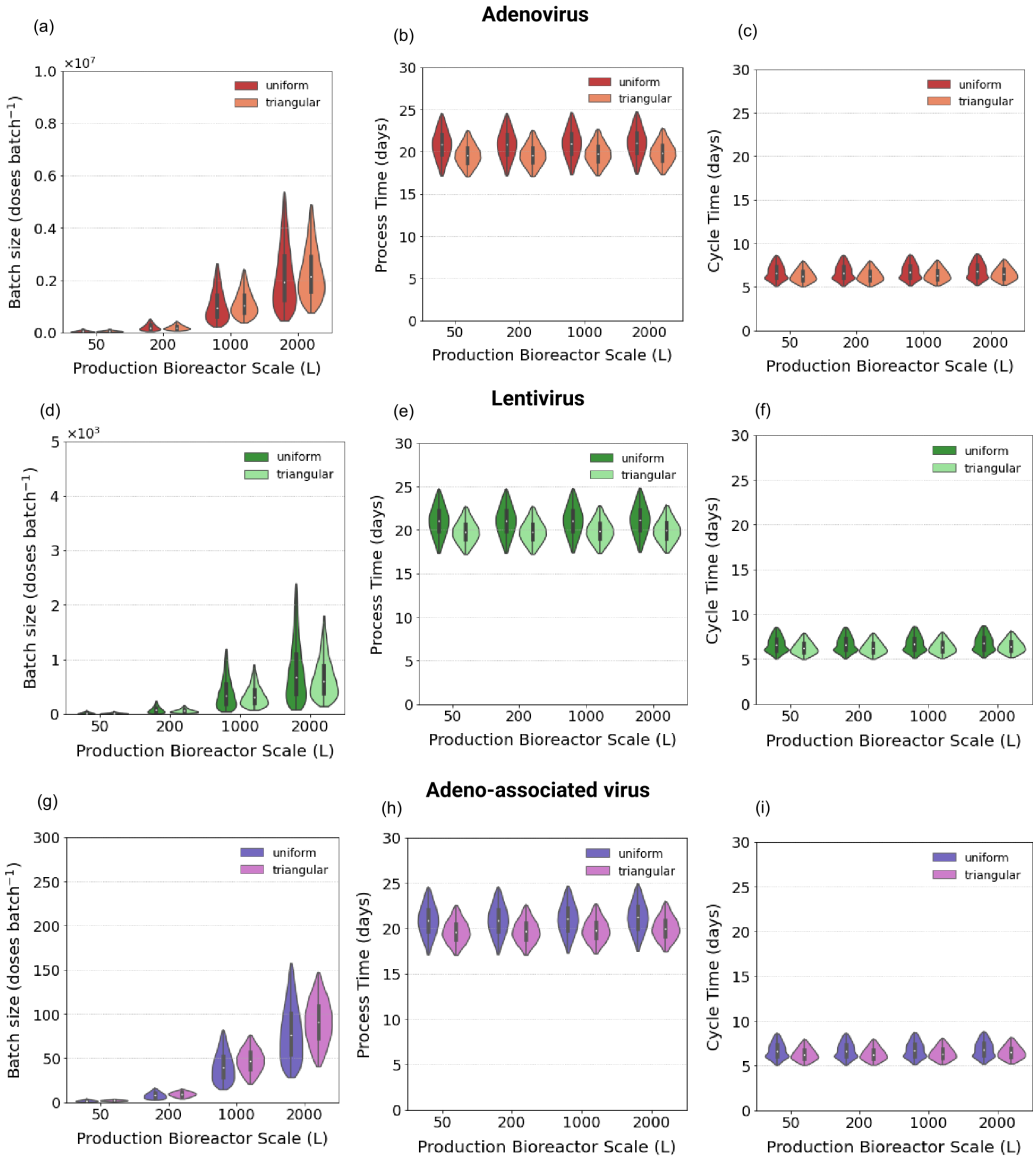


Figure 3.4: Violins plots for throughput related quantities comparing two input distributions including uniform and triangular. Scale-dependent results are shown for (a) process time for AdV, (b) process time for LV, (c) process time AAV, (e) cycle time for AdV, (e) cycle time for LV, (f) cycle time for AAV, (g) batch size for AdV, (h) batch size for LV, (i) batch size for AAV.

production scale increases (Figure 3.4). This is due to the assumption that titer ranges do not vary across scales and directly multiply main bioreactor working volume which increases proportionally with scale. Additionally, the criticality of USP titer for batch size variability identified via GSA explains the prevalence of this trend over the impact of DSP recoveries, which correspond to a fractional change in batch size and would contribute to a linear increase instead.

Batch sizes for AdV-based vaccines applications consistently remain larger than LV-based and AAV-based gene therapies. Specifically, when considering a uniform distribution, AdV batch sizes range at 2000L around medians of 1.9×10^6 doses/batch (IQR=1.2–2.9×10^6 doses/batch), decreasing to a median of 9.4×10^5 doses/batch (IQR=5.9×10^5–1.4×10^6 doses/batch) at 1000L scale, up to a median of 4.7×10^4 (IQR=2.9×10^4–7.2×10^4 doses/batch) at 50L (Figure 3.4a). Compared to AdV, LV batch sizes remain lower ranging from around 672 doses/batch (IQR=354–1133 doses/batch) at 2000L to batch sizes of 17 doses (IQR=9–28 doses/batch) at 50L (Figure 7B). Median values for AAV batches at 2000L correspond to 76 doses/batch (IQR= =53–102 doses/batch) and drop to significantly lower amounts at 50L, with 2 doses/batch (IQR=1–3 doses/batch) (Figure 3.4).

If a triangular distribution is assumed, tighter ranges of batch sizes are observed for each scale. The assumed most probable value for USP titers and recoveries is found to impact the shape of the distribution. In the LV process, most probable input values for USP titer and DSP recoveries are generated at a midpoint within the considered range and this translates in a symmetrically skewed distribution compared to the uniform case, with median values preserved for all scales and tighter IQR ranges. In the case of AdV and AAV, USP peak values correspond to the upper bound whereas IEX and affinity chromatography yields correspond to the lower bound. As USP titer is identified as most impactful to this output, median values for AdV and AAV batch sizes increase for all scales while IQR ranges decrease, as more USP titer input samples are generated in the upper end of the specified ranges.

Process times remain almost unchanged across scales (Figure 3.4c, f and i). This is due to the assumption that the number of expansion steps and task lengths are independent of scale. Uniform distributions result in medians of process times around 21.2 days (IQR=19.9–22.5 days) while a median cycle time of 6.8 days (IQR = 6.1–7.6 days) is found for all processes. When triangular

distributions are assumed, more input samples are generated around the most probable value of 3.6 days for cell expansion tasks and 2 days for viral growth. Output samples result distributed around lower median values and a tighter IQR range is recorded, with medians of 19.8 days (IQR = 18.9–20.8 days) for process time and 6.1 days (IQR = 5.8–6.9 days) for cycle time.

3.3.2.2 Cost outputs

Figure 3.5 displays the results of the UA for all processes and scales with a focus on uncertainty distributions for the cost outputs. Smaller scales result in lower capital investments as equipment of smaller volume is required in each process.

AdV capital costs drop from a median of 242.7×10^6 \$ (IQR = 242.4–243.0×10^6 \$) at 2000L to 80.5×10^6 \$ (IQR = 80.5–80.6×10^6 \$) at 50L scale (Figure 3.5a). Overall, the process requires the highest capital expenditure compared to LV and AAV due to the F&F process sections at all scales. LV capital costs range from 128.5×10^6 \$ (IQR = 128.2–128.9 $\times10^6$ \$) at 2000L to costs of 71.3×10^6 \$ (IQR = 71.2–71.4×10^6 \$) at 50L scale (Figure 3.5b), whereas in AAV manufacturing at 50L capital costs remain higher than LV, resulting in a median of 91.6×10^6 \$ (IQR = 91.5–91.7×10^6 \$) (Figure 3.5c). Uncertainties around the capital costs remain narrow at all scales, with a variability introduced via the working capital contribution previously characterised via GSA.

Scaling down batch sizes results in lower operating expenditures per batch for all vectors. Production costs for AdV batches remain higher than AAV batches and LV batches at all scales. For instance, in the uniform distribution case AdV production costs decrease from 1.41×10^6 \$/batch (IQR = 1.37–1.45×10^6 \$/batch) at 2000L, to costs at 1000L at medians of 0.90×10^6 \$/batch (IQR = 0.87–0.93×10^6 \$/batch) up to 0.36×10^6 \$/batch (IQR = 0.33–0.38×10^6 \$/batch) at 50L (Figure 3.5a). These ranges remain higher than the LV and AAV process scales, although processes converge to an almost equivalent expenditure of medians 0.32–0.39×10^6 \$/batch at 50L (Figure 3.5d and g). Additionally, assuming triangularly distributed resource costs with central peak values results in a symmetrical tightening of IQR ranges around preserved median values. OpEx per batch is strongly dependent on resource costs realizations as observed in the GSA analyses and the distribution pattern is preserved. For both distributions larger IQR are observed with increasing scale due to the larger amount of raw

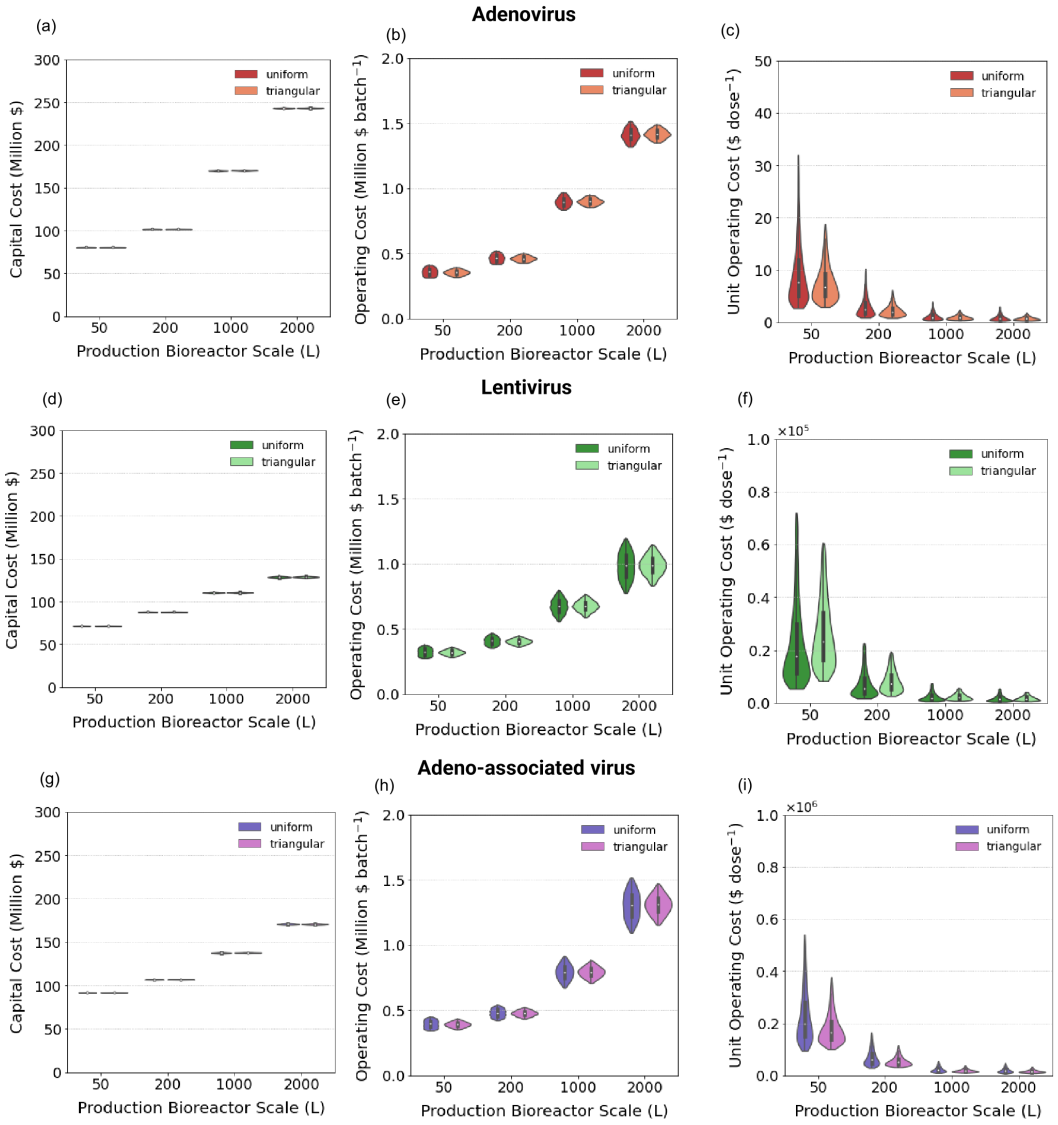


Figure 3.5: Violins plots for cost related quantities comparing two input distributions including uniform and triangular. Scale-dependent results are shown for (a) Capital cost for adenoviral (AdV), (b) capital cost for lentiviral (LV), (c) capital cost for adeno-associated vector (AAV), (d) operating cost for AdV, (e) operating cost for LV, (f) operating cost for AAV, (g) unit cost for AdV, (h) unit cost for LV, (i) unit cost for AAV.

material required to manufacture larger batch sizes.

Unit production costs result higher at lower scales due to economies of scale for all processes, with a 10–20-fold increase at 50L compared to 2000L (Figure 3.5c, f and g). If OpEx per batch for each vector is at similar orders of magnitude, converging to an equivalent range for smaller scale, differences across vectors are significant when costs are normalized by batch size ranges. Specifically, for the uniform distribution case, a 10-fold increase in unit cost for AdV-based vaccines translates in an increase from 0.74 $/dose (IQR = 0.48–1.16 $/dose) at 2000L to 7.56 $/dose (IQR = 4.90–12.2 $/dose) at 50L (Figure 3.5f). For LV and AAV products, manufacturing at lower scale results in sharper increases in unit costs, as batch sizes are found to be lower by orders of magnitude compared to AdV vaccines. At 50L scale, LV unit production costs range at 19127 $/dose (IQR = 11396–36640 $/dose) compared to a median of 1459 $/dose (IQR = 866–2730 $/dose) at 2000L scale (Figure 3.5e). Unit AAV costs at the lowest 50L scale reach a median of 197973 $/dose (IQR = 146360–283598 $/dose) (Figure 3.5g). The impact of triangular distributions is also assessed and is related to the trends observed for the batch size for each vector application. These analyses highlight that manufacturing at larger scales is expected to lead to significant cost reductions, specifically in the case of *ex vivo* LV and *in vivo* AAV-based products which are characterized by smaller batch sizes and typically higher dose sizes.

3.3.3 Uncertainty characterisation

3.3.3.1 Throughput outputs

The results of GSAs with respect to batch size, batch time and cycle time are displayed in Figure 3.6. First-order sensitivity indices highlight that viral vector batch size variability is primarily caused by the titer input variability, followed by the performance of the chromatography steps.

Titers and working volumes vary within orders of magnitude up to $O(10^2)$ and are expected to result in larger variance in the batch size, as opposed to DSP recoveries which translate into a 0–1 fractional change of the batch size. This is seen in the AdV process, where the USP titer, working volume and IEX yield have respective first-order contributions Si of 0.50, 0.35, and 0.06. Significant combined effects between titer and working volume are observed ($S_{ij} = 0.06$), as

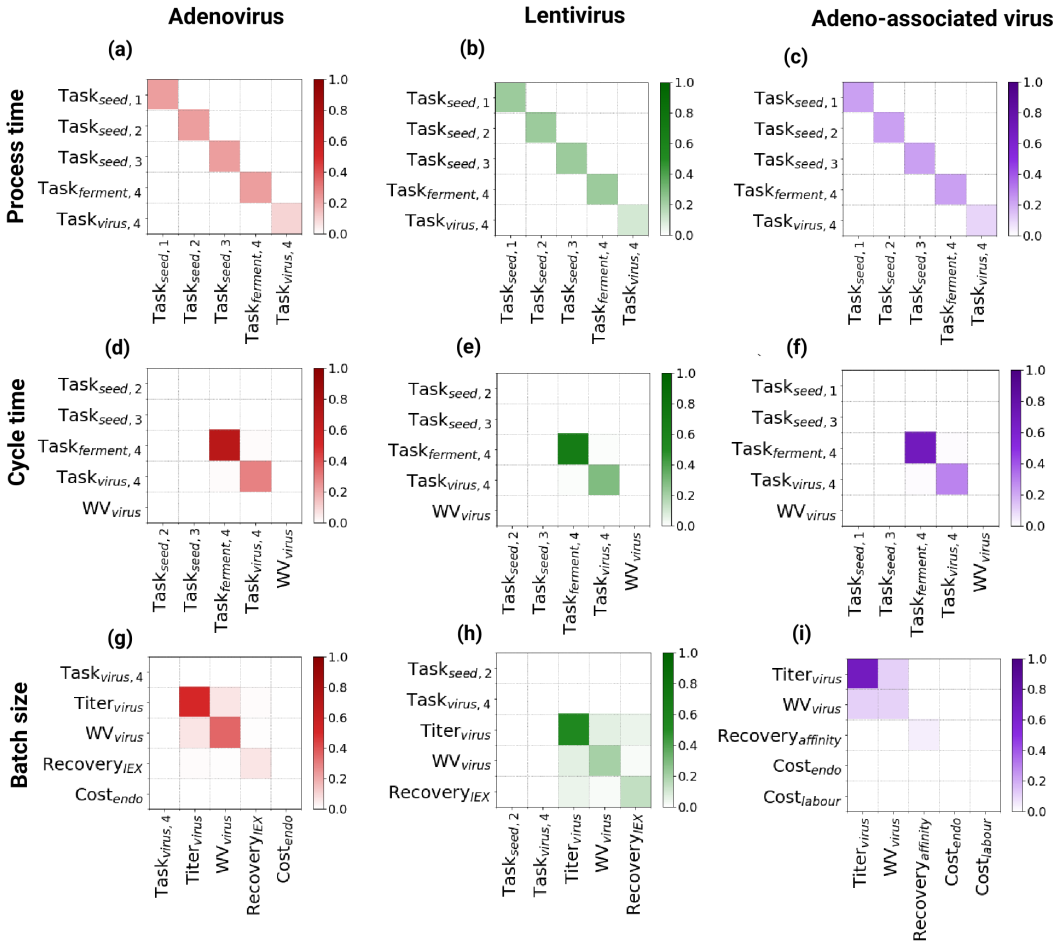


Figure 3.6: Global sensitivity analysis (GSA) plots for throughput-related outputs of each viral vector platforms and highest ranking inputs. Each plot displays first-order sensitivity indices S_i along the diagonal and second-order indices S_{ij} in squares combining two inputs: (a) process time AdV, (b) process time for LV, (c) process time for AAV, (e) cycle time for AdV, (e) cycle time for LV, (f) cycle time for AAV, (g) batch size for AdV, (h) batch size for LV, (i) batch size for AAV.

well as weaker interactions ($S_{ij} = 0.01$) with the IEX yield for both inputs.

Similar trends are observed in the LV process, with first-order effects for the USP titer ($S_i = 0.53$), working volume ($S_i = 0.20$) and IEX yield ($S_i = 0.14$) on batch size variability. Combined effects between USP titer and working volumes ($S_{ij} = 0.03$) are observed alongside those with the IEX step ($S_{ij} = 0.01$).

Finally in AAV manufacturing, first-order effects are once more observed for inputs related to USP productivity, namely USP titer ($S_i = 0.68$) and working volume ($S_i = 0.11$). The recovery of the affinity chromatography step stands out as impactful to batch size ($S_i = 0.05$), stronger than the IEX step ($S_i = 0.02$), which was in fact excluded at the GSA I step (Appendix B2-3). Despite sampling over an equivalent range of chromatography yields, the fractional change in the IEX is computed on a smaller batch size, which results in lower contributions to the computed batch size variance. A strong combined effect on the output ($S_{ij} = 0.11$) is observed for USP titer and working volume, while interactions with chromatography yields remain lower ($S_{ij} < 0.05$).

Filtration steps such as UF/DF do not impact the batch size significantly (Appendix B). These purification steps tend to be rather standardized in the industry, with recoveries that do not go below 90% for both AdV and AAV. In the case of LV instead, filtration recoveries can reach 70% as reported by Perry and Rayat (2021). As a result the variability introduced by titer and chromatography step yields is remarkably more significant.

The length of time to produce the batch solely depends on the assumed task lengths (Figure 3.6a-c). As processes are assumed to not be vector specific until transfection/infection and the longest tasks take place in the USP section, variability in process time is mainly attributed to task durations in the seed cell culture train (Steps 1–3) and the main bioreactor (Step 4), followed by viral growth. This is confirmed by the value of sensitivity indices ($S_i = 0.22$) for all cell tasks for all processes and the viral vector growth with a lower index ($S_i = 0.099$).

The cycle time for each of the processes is impacted by the cell culture step duration in the main bioreactor ($S_i = 0.69$) and the transfection task length ($S_i = 0.28$) which constitute the bottleneck of the process (Figure 3.6d-f).

3.3.3.2 Cost outputs

A further set of GSA plots is presented in Figure 3.7, which illustrate how the combination of critical operational inputs and resource cost inputs impact cost-related outputs, including capital costs, operating costs per batch and operating cost per dose.

Working capital includes the upfront cost of materials for the initial 30 days of operation. This cost would be expected to vary according to the cost of the resources consumed for each batch and the actual number of batches completed within the 30-day period. Therefore, pDNA cost is the main contributor to working capital variability in the LV process ($S_i = 0.57$) and the AAV process ($S_i = 0.47$) followed by contributions from other resources ($S_i = 0.010.09$) (Figure 3.7a-c). Weaker first-order contributions from inputs that impact bottleneck lengths and number of batches is also observed, namely the task lengths relate to cell expansion and viral propagation in the main bioreactor ($S_i = 0.080.19$). In the case of AdV, first-order effects for these task lengths ($S_i = 0.52$, $S_i = 0.21$) prevail over materials resource costs (Figure 3.7a).

The variability in OpEx per batch for each vector is primarily introduced by materials and operating costs (Figure 3.7d-f). Given the inputs considered and the equipment size remaining fixed, the facility-dependent component of the cost does not contribute to the observed output variability. Plasmid DNA cost is the most prominent in LV ($S_i = 0.89$) and AAV manufacturing ($S_i = 0.75$) as it contributes to the largest share of the cost (Figure 3.7e and f). The OpEx per batch also varies due to the cost of endonuclease, with first-order contributions in both the LV process ($S_i = 0.09$) and the AAV process ($S_i = 0.09$). A range of weaker first-order sensitivities ($S_i = [0.080.02]$) is also observed with respect to labour and medium cost in both LV and AAV manufacturing.

GSA for the AdV process highlights stronger first-order effects for endonuclease costs ($S_i = 0.41$) and the labour rate ($S_i = 0.32$), as pDNA is not consumed by the process (Figure 3.7d). Additionally, the amount of medium consumed and its cost show significant first-order ($S_i = 0.14$) and weak second-order effects ($S_{ij} = 0.01$), as they impact the total cost of medium consumed which is factored in the materials cost. This second order effect becomes insignificant in LV and AAV manufacturing given the stronger Si of pDNA cost.

Focusing on cost per dose in AdV manufacturing, first-order effects from USP

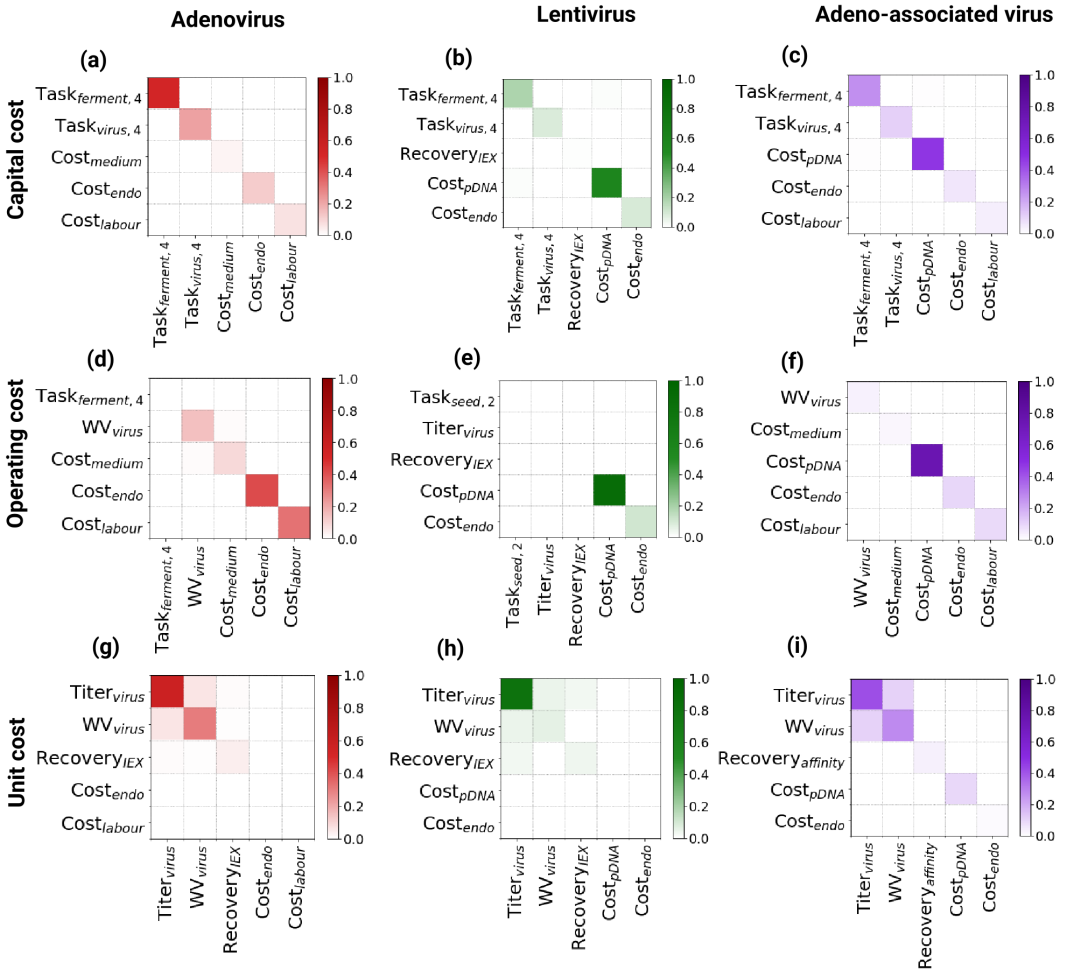


Figure 3.7: Global sensitivity analysis (GSA) plots for cost-related outputs of each viral vector platforms displaying highest ranking inputs. Each plot displays first-order sensitivity indices S_i along the diagonal and second-order indices S_{ij} in squares combining two inputs: (a) capital cost for AdV, (b) capital cost for LV, (c) capital cost for AAV, (d) operating cost for AdV, (e) operating cost for LV, (f) operating cost for AAV, (g) unit cost for AdV, (h) unit cost for LV, (i) unit cost for AAV.

titer ($S_i = 0.57$), working volume ($S_i = 0.30$) and IEX yield ($S_i = 0.042$) are recorded (Figure 3.7g). Similarly in LV manufacturing, significant inputs include USP titer ($S_i = 0.80$), working volume ($S_i = 0.60$) and IEX yield ($S_i = 0.036$) (Figure 3.7h). In AAV manufacturing, the impact of pDNA cost is also significant ($S_i = 0.09$), alongside the expected USP titer ($S_i = 0.41$), working volume ($S_i = 0.27$) and affinity chromatography yield ($S_i = 0.036$) (Figure 3.7o). This would be explained by the strong first-order contribution found for the OpEx per batch in AAV manufacturing, based on which the unit cost is computed. Second-order effects between USP titer, working volume and chromatography yields also propagate to the unit production cost, given its dependence on batch size.

These results highlight the importance of either improving the productivity per batch and/or lowering dose sizes. It is in fact expected that lowering doses for gene therapy LV and AAV products especially would lower costs of production per unit by orders of magnitude. Additionally, reducing pDNA requirements for transfection would contribute to lowering AAV unit costs production.

3.3.4 Supply chain implications

The quantification of scale-dependent process uncertainty is a crucial step to assist decision-making throughout scale-up. A staged approach to scale-up mitigates financial risks of investment, whilst reducing possible benefits from economies of scale in early phases of process development.

In Phase I clinical trials, manufacturing is often allocated to existing small scale assets. As clinical trials progress and product demand increases manufacturers assess the maximum scalability of available resources before switching to larger scales. In this space, the probability distributions and computed IQR ranges can be used as means to quantify worst-case and best-case performance of each KPI considered across scales. In addition, the shape of the probability distribution displays the likelihood of each scenario and can help quantify risks. Furthermore, the critical parameters identified via the GSA inform manufacturers on the key driver that would help achieve improved performance with respect to each KPI.

For instance, AAV manufacturers can assess the scalability of a 200L production platform, which manufactures an approximate number of 10 doses per batch

(Figure 3.4g). A higher value may be achieved with combined improvements of USP titers, maximising working volumes and improved affinity chromatography, as identified via GSA. Unit production costs are directly impacted by batch size; hence, in the best case scenario these would be lowered to approximately 44,886 $/dose. With a switch to larger scales, a 1000L reactor would confidently ensure 28-53 doses at 14,874-28,900 $/doses, with expected further reductions at 2000L scale (Figure 3.5i). In the AAV case the uncertainty distributions are found to not overlap for all IQR ranges, which suggest that one can more confidently differentiate performance and benefits across scales. This type of assessment could assist strategies in prioritising investments in larger scale for high-dose products and quantify substantial benefits of switching to larger scales early on.

In the LV case, one could infer that the manufacturing uncertanties may result in more complex trade-offs during scale-up. Batch size outputs and IQR ranges overlap across scales which challenges the differentiation of scales Figure 3.4d. For instance, a 1000L manufacturing process ensures batch sizes of around 336 doses per batch. If production is allocated to this asset, manufacturers may consider that the potential of process improvements and conclude that the benefit of switching to larger scales is less straightforward. With scale up, a 2000L process can supply batches as low as 354 doses/batch correspoinding to costs of 2,730 $/dose. In this case, even if process knowledge is integrated, the triangular distribution outputs correspond to similar cost. This highlights the importance of considering drops in yields occuring during scale-up which may result in reverse economies of scale. Of course, if scenario probabilities are factored in, larger scales become the most reliable option.

Uncertainty analysis is crucial in scenarios that require rapid scale-up and deployment of capacity. In the context of AdV-based vaccines, manufacturers may want to install capacity which has the largest probability to meet global demands faster. As shown in Figure 3.4a, processes operating at 2000L scale have a larger probability of supplying larger amounts compared to 1000L. However, if unit cost uncertainty is considered (Figure 3.5c), one can infer that these correspond to nearly equal ranges. This makes both scales cost-competitive, which highlights scale-out through staggering or distributed manufacturing networks is a good alternative to scale-up. In this fashion, network cycle times and transport delays could also be reduced.

Overall, the reduction of network cycle time is a pressing need for production

platforms that have to readily supply to a demand which may present short-term fluctations. In the case of LV for example, manufacturers of C&GT therapy products producing LV in house may need to carefully assess the benefits of reduction in cycle time with process intensification versus staggering equipment, thus increasing equipment cost and sacrificing cost-effectiveness.

3.4 Conclusions

In this Chapter, we have presented a framework for the characterisation and quantification of product specific and scale-dependent uncertainty. This is a crucial first step that underpins the development of decision-making tools that optimise capacity planning decisions with an underlying knowledge of process performance uncertainty.

For each process, we are able to identify and quantify key input uncertainties that contribute to manufacturing KPIs of batch size, batch time, bottleneck time and costs. This helps understand key input-output correlations. By sampling the input space through Sobol'-based quasi random sequences we are able to construct uncertainty ranges for the output KPIs of interest which account for potential interactions between input parameters and achieve good coverage of the input uncertainty space. Through this approach different input distributions can be tested.

This exercise quantifies uncertainty ranges of performance which assume no underlying correlations. As a result, decision-makers can directly use the comparative plots to quantify benefits of scale-up and scale-out approaches for each product type and develop manufacturing strategies accordingly, comprising single- or multi-product production. Alternatively, the data set constructed through this approach can be used systematically in optimisation-based planning frameworks. In Chapter 4 we assess the impact of uncertainty integration on optimal network design. Furthermore, in Chapter 5 we use the acquired knowledge of the uncertainty space to quantify the actual flexibility of a given design. Finally, in Chapter 6 we consider the realisation of process uncertainty alongside demand uncertainty in multi-period staged investment planning.

Chapter 4

Integration of process uncertainty

Outputs of this chapter have been disseminated in:

- **Sarkis, M.**, Tak, K., Chachuat, B., Shah, N. & Papathanasiou, M.M. (2021). Towards resilience in next-generation vaccines and therapeutics supply chains, *Book of Abstracts of the 32nd European Symposium on Computer Aided Process Engineering,* 931-936.
- **Sarkis, M.**, Shah, N. & Papathanasiou, M.M. Resilient pharmaceutical supply chain planning: assessment of strategies for process uncertainty integration in network design problems. *Submitted to Computers & Chemical Engineering*, Under review.

4.1 Introduction

With the continued growth of novel biotechnology driven drugs and an increased need for pandemic preparedness, the advanced therapeutics sector requires decision-support tools for informed investment planning. These tools should aim at helping to mitigate shortages risks and assessing the impact of manufacturing uncertainties on attainable production targets and required investments. This becomes ever important as the sector transitions to flexible and modular multi-product manufacturing facilities.

In recent years, critical shortages in the viral vector sector (Gardner B, 2020) were reported, resulting for a need to scale-up to cater for different markets, while manufacturing platforms were still under development (Capra et al., 2022; Perry and Rayat, 2021; Merten et al., 2011; Ansorge et al., 2009; Bauler et al., 2020a). This demonstrated a need to account for process uncertainty in planning initial investments and production campaigns.

In this space, computer-aided tools have an inherent potential to assist decision-makers throughout the early stages of scale-up when operating under significant process uncertainty. Modelling and optimisation of supply chain networks can help understand the impact of various sources of uncertainty on network performance and quantify the benefits of systematically integrating uncertainty to yield *uncertainty-aware* optimisation solutions.

The first step towards building tools for planning under uncertainty consists of the quantification of uncertain parameters. Chapter 3 proposes a methodology for the quantification of process specific uncertainty in the context of emerging pharmaceutical products. The second step consists in formulating an optimisation problem which identifies good supply chain network designs that account for the quantified uncertainty.

One approach to optimisation under uncertainty is robust optimisation, which is considered a *risk-averse* strategy. By seeking solutions that are feasible for all realisations of the uncertainty, the approach is suited for problems where recourse actions and information of uncertainty distributions is not available (Janak et al., 2007; Lappas and Gounaris, 2018; Wiebe et al., 2019). In strategic planning problems, this may lead to overly-conservative solutions. Hence, long- and mid-term planning problems under process and demand uncertainty are often modelled as multi-stage stochastic programms with recourse actions.

It is worth noting that the quantified uncertainty ranges represent a two-stage uncertainty. At some instant in the future time horizon, one would think that knowledge real process capability will be acquired. A number of efforts in the literature have focused on multi-stage stochastic capacity planning under future demand uncertainty, as this is considered to realise over time and influenced by outcomes in the previous time periods (Gatica et al., 2003a; Levis and Papageorgiou, 2004; Tsang et al., 2007a; Colvin and Maravelias, 2008). Zeballos et al. (2014) combine demand uncertainty and process uncertainty in a multi-stage scenario tree, thus accounting for a recurring uncertainty in process outcomes at each stage.

Effectively, these multi-stage uncertainty modelling approaches do not capture the potential two-stage nature of process uncertainty in the Advanced Theray Medicinal Products (ATMPs) context quantified through Chapter 3. The work by Zeballos et al. (2014) models strategic process uncertainty, recurring over the time horizon and considers time-dependent decisions. As a first step to formulate an analogous optimisation for the process uncertainty quantified for ATMPs, we consider the time-independent supply chain design problem under process uncertainty through a two-stage stochastic optimisation. Thus, in this Chapter we develop uncertainty integration strategies for the quantified scale-dependent uncertainty ranges.

A crucial trade-off between computational cost and the accuracy of the scenario tree emerges in stochastic formulations. Several scenario-tree generation strategies, including moment-matching methods and scenario-reduction techniques Kaut (2021) are reported in the literature. In this Chapter, we propose a sampling method which generates scenarios based on the empirical samples obtained in Chapter 3. In this fashion, we preserve the assumed independence of each manufacturing process as well as the joint probabilities of the parameters with respect to each individual process.

Through this approach to stochastic programming we explore the impact of the number of representative samples used as representative scenarios in the optimisation, with the aim of minimising computational costs. Following this, we test the *out-of-sample* stability of each solution through a stochastic simulation sampling from the original datasets.

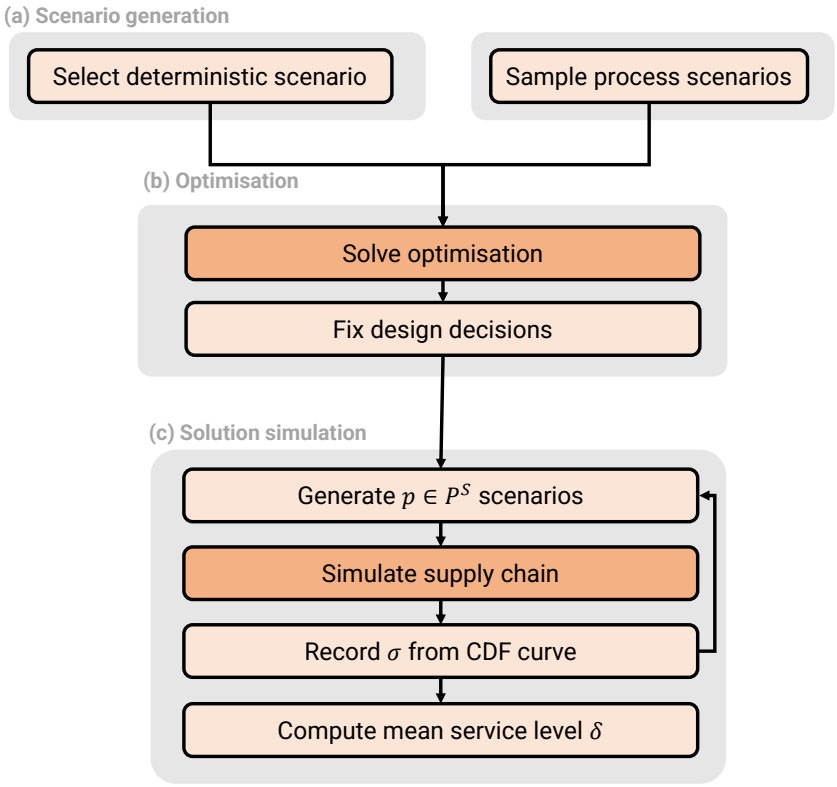


Figure 4.1: Optimisation-simulation framework: (a) process scenarios are generated for the optimisation (b) the optimisation is formulated as a deterministic problem or a stochastic program, (c) the simulation scheme is implemented to quantify the statistical performance of each optimisation solution.

4.2 Methodology

The framework presented in this study consists of (i) a process scenario generation step, (ii) an optimisation step for network design and a (iii) simulation step for statistical performance quantification (Figure 4.1). The following sections illustrate the methodology adopted in the development of the optimisation-simulation framework and discuss the case studies used to illustrate the capabilities of the developed tool.

4.2.1 Problem statement

The problem considered in this study focuses on capacity planning and is based on the network superstructure illustrated in Figure 4.2. The objective of the planning problem is to minimise total supply chain capital and operating costs

and determine the following decisions, including scenario-independent first stage decisions and scenario-dependent second-stage decisions (recourse actions):

- a network structure and links between nodes;
- investment in manufacturing and storage facilities;
- scenario-dependent scales of manufacturing and number of parallel lines;
- scenario-dependent production levels at each manufacturing node;
- scenario-dependent transportation flows between existing nodes;
- scenario-dependent supplied quantities at each demand node.

To do this, the optimisation considers a range of inputs including:

1. a set N of nodes n;
2. a set A of manufacturing scales a;
3. a set S of manufacturing process sections s;
4. a set L^S of manufacturing process lines l;
5. a set P of manufacturing process scenarios p;
6. datasets if scale-dependent batch sizes, batch times and costs;
7. data on distances between each node n;
8. data on transportation costs;
9. data on storage costs;
10. demand of a single product in demand zones $n \in N^Z$.

The challenge in this capacity planning problem is the uncertainty in techno-economic performance of the manufacturing nodes $n \in N^J$. As found in Chapter 3, underlying manufacturing uncertainties include batch size, process times, bottleneck lengths and costs and are scale-dependent and product dependent. In this work, we investigate the benefit of integrating this quantified uncertainty within the investment planning activity. We consider different optimisation strategies, namely:

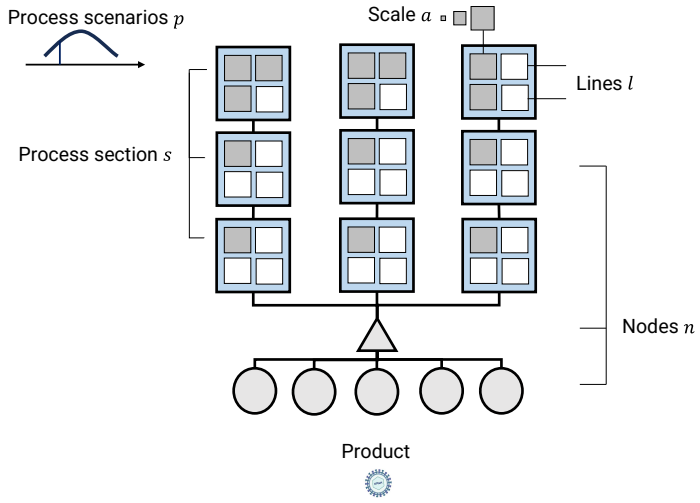


Figure 4.2: Superstructure for optimisation of single-product (LV) case study. Nodes n comprise manufacturing nodes (square), storage (triangle) and demand zones (circles). Each manufacturing node is modelled as a set of process sections s comprising parallel manufacturing lines l at scale a. Process scenarios p are generated as input datasets for the optimisation-simulation framework.

1. *Best-case*: deterministic optimisation with parameters held at best-case value;

2. *Median-case*: deterministic optimisation with parameters held at median value;

3. *Worst-case*: deterministic optimisation with parameters held at worst-case value;

4. *Stochastic (quasi-random)*: stochastic optimisation using scenarios generated via Sobol' sampling.

5. *Stochastic (random)*: stochastic optimisation using scenarios generated via random sampling.

4.2.2 Process scenario generation

The first step within the framework (Figure 4.1) is the integration of process scenarios as inputs for the planning problem. Specifically, the aim of this step is the assignment of values to the vector of process inputs ξ, comprising batch

sizes χ, process time α, bottleneck length τ (i.e. time between two consecutive batches) and cost parameters c.

4.2.2.1 Input dataset structure

In Chapter 3, a total of $\Omega = 1000$ samples are generated through the uncertainty analysis (UA) framework. Samples are available for each scale a of a product of interest and each sample corresponds to a unique value of each key performance indicator (KPI), denoted here as ξ. It is crucial to note that each uncertainty analysis is conducted separately, which yields a $d-$dimensional uncertainty distribution of $d = A$.

We define a matrix of input data samples from all uncertainty analyses as Θ, with dimensions $\Omega \times d$. Each sample in the Θ matrix maps to a matrix of values Γ for each parameter ξ, namely parameters of batch sizes α, process time τ and cost parameter c.

In the deterministic optimisation cases, the value of the parameters ξ can be assigned by finding the best, median and worst case scenario respectively for each KPI. In this case, $p = \{P_1\}$ and the parameters are set to unique values retrieved from the corresponding Γ for the KPI ξ directly.

In the stochastic optimisation cases, the problem should integrate a sufficient number of scenarios to accurately represent the uncertainty space. First, we generate a sub-matrix of samples θ, which embeds the first $\omega = 128$ rows of Θ. This is justified by the fact that all uncertainty analyses were conducted via Sobol-based sampling and the shape of the uncertainty distributions for all KPIs is found to be preserved for a subset of samples. This is because of the systematic generation of low discrepancy sequences, which ensure good coverage of the uncertainty space from relatively low number of samples. The dimensionality reduction is expected to help achieve a better coverage of the uncertainty space with a smaller number of random or quasi-random samples for stochastic programming.

Secondly, given a matrix of parameter samples θ, each column is randomised to address a potential bias arising from the Sobol-based uncertainty analyses used to generate the data. Whilst this step is not strictly required for random sampling given the inherent randomness of the approach, it is crucial for the quasi-random sampling counterpart. In fact, it helps decouple the effective impact of increasing

a number of scenarios for optimisation from pattern-based bias.

4.2.2.2 Sampling methodology

Given the randomised θ matrix of $\omega \times d$ dimension, the generation of optimisation scenarios for stochastic programming leverages on sampling with respect to θ. Specifically, we define P^O as the set of optimisation scenarios for the stochastic optimisation, which is made of O elements.

In the random sampling-based stochastic programming a matrix X_{random} of $O \times d$ integers between 1 and ω is generated. This ensures the random integers correspond to elements within the submatrix θ.

$$X_{\text{random}} \in \{1, 2, \ldots, \omega\}^{O \times d} \tag{4.1}$$

The matrix X_{random} can be used to set the value for the general KPI parameters ξ:

$$\xi_{pd} = f(X_{\text{random}}, \theta, \Gamma_\xi) \quad \forall p \in P^O, d \tag{4.2}$$

Where f can be understood as a mapping function retrieve values from θ based on X_{random} and sets KPIs values ξ based on the corresponding Γ. This means assigning values for batch size χ, process time α, bottleneck length τ and cost parameters c for each p and dimension d.

Similarly, in the quasi-random sampling-based stochastic program a matrix X_{Sobol} of $O \times d$ integers between 1 and ω is generated.

$$X_{\text{Sobol}} \in \{1, 2, \ldots, \omega\}^{O \times d} \tag{4.3}$$

The matrix of integers X_{Sobol} generated based on low-discrepancy Sobol' sequences, θ are used to assign the respective values for the general ξ parameters based on the corresponding value matrix Γ for the parameter:

$$\xi_{pd} = f(X_{\text{Sobol}}, \theta, \Gamma_\xi) \quad \forall p \in P^O, d \tag{4.4}$$

4.2.3 Optimisation formulation

The optimisation step in Figure 4.1 considers a single-period single-product planning problem for the manufacturing of an LV-based product for gene therapy

applications. The optimisation is formulated as a mixed integer linear problem (MILP). Deterministic formulations can be obtained by considering $P = \{P_1\}$, thereby inputting a single scenario realisation for the uncertain parameters. In the stochastic case, the N samples drawn to generate the uncertain process realisations require $P = \{P_1...P_N\}$, based on the sampling procedures described in Section 4.2.2.

4.2.3.1 Objective function

The objective of the optimisation is to minimise the total expected cost z of the supply chain (Eq.4.5). Total costs include the scenario-dependent capital costs and operating costs of manufacturing, storage costs and transport costs, multiplied by the respective probability ϕ_p.

$$\min \; z = \sum_{p \in P} \phi_p (TC_p^{CAPj} + TC_p^{OPj} + TC_p^{OPk} + TC_p^{T}) \tag{4.5}$$

Where TC^{CAPj} is the total capital manufacturing cost, TC^{OPj} is the total operating manufacturing cost, TC^{OPk} is the total operating storage cost and TC^{T} is the total transport. These are all scenario-dependent quantities as they are functions of scenario-dependent process realisations p for throughput and cost-related quantities. It is worth highlighting that the investment decisions are scenario-independent, whereas the cost parameters are uncertain and therefore scenario-dependent.

In deterministic cases, one single scenario P_1 is considered of probability $\phi_{P_1} = 1$. In stochastic cases, all scenarios are assumed to be equiprobable, given that the original samples generated through UA in Chapter 3 are identically distributed samples (IID) and a subset of these is selected. Hence the probability ϕ_p is computed as:

$$\phi_p = \frac{1}{O} \quad \forall p \tag{4.6}$$

4.2.3.2 Capital manufacturing costs

Capital costs of manufacturing (Eq.4.7) are computed as a sum of facility installation costs and product-specific equipment investment costs.

$$TC_p^{CAPj} = \sum_{n \in N^J} C_{np}^{EQj} \quad \forall p \tag{4.7}$$

The term in the sum of Eq. 4.7 is C^{EQj}, which is a scenario-dependent variable accounting for the total cost of equipment investment in node $n \in N^J$ and process scenario p.

$$C_{np}^{EQj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} c_{asip}^{EQjl} M_{nasl} \quad \forall n \in N^J, p \tag{4.8}$$

The equipment installation cost C^{EQj} is computed as a function of (i) *first-stage* binaries M, which mark whether equipment for the manufacturing at scale a is available and installed in lines l of section s in plant $n \in N^J$, (ii) scenario-dependent capital cost realisations for parameters c^{EQjl}(Eq.4.8).

4.2.3.3 Operating manufacturing costs

Total operating costs TC^{OPj} (Eq. 4.9) of manufacturing correspond to a summation of annual variable costs C^{VARj}, facility related costs C^{FACj} and labour costs C^{LABj} for each manufacturing plant $n \in N^J$, process realisation p at time t.

$$TC_p^{OPj} = \sum_{t \in T} \sum_{n \in N^J} C_{tnp}^{LABj} + C_{tnp}^{VARj} + C_{tnp}^{FACj} \quad \forall p \tag{4.9}$$

Labour costs C^{LABj} for each time period t, manufacturing plant $n \in N^J$ and scenario p are computed (Eq. 4.10) as a summation of the (i) labour costs c^{LABjl} parameters for each scale a, process section l, product i and process realisation p multiplied by (ii) batches B respectively which vary with respect to time t, plant $n \in N^J$, scale a, process section s, production line l, product i and process realisation p:

$$C_{np}^{LABj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} c_{asp}^{LABjl} B_{naslp} \quad \forall n \in N^J, p \tag{4.10}$$

Variable costs C^{VARj} for each manufacturing plant $n \in N^J$ and scenario p are computed (Eq. 4.11) as a multiplication of the (i) variable costs parameters c^{VARjl} for each scale a, process section s and process realisation p by (ii) batches B respectively which vary with respect to plant $n \in N^J$, scale a, process section s, production line l and process realisation p:

$$C_{np}^{VARj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} c_{asp}^{VARjl} B_{naslp} \quad \forall n \in N^J, p \tag{4.11}$$

Facility-dependent costs C^{FACj} are tied to depreciation and equipment maintenance costs. In Eq. 4.12, facility-dependent costs for every manufacturing plant $n \in N^J$ and process realisation p are computed as multiplication of (i) facility costs parameters c^{FACjl} for each scale a, process section l and process realisation p by (ii) equipment installation-related binaries M respectively which vary with respect to plant j, scale a, process section s and production line l.

$$C_{np}^{FACj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} \sum_{i \in I} c_{asp}^{FACjl} M_{nals} \quad \forall n \in N^J, p \tag{4.12}$$

4.2.3.4 Logistics-related costs

Total operating costs of storage TC^{OPk} are a function of (i) the batches handled in the process Q, which corresponds to the number of batches of scale a shipped from upstream storage node $n \in N^K$ to downstream demand zone $m \in N^Z$ in scenario p, (ii) the cost parameter c^{OPk} at scale a.

$$TC_p^{OPk} = \sum_{n \in N^K} \sum_{m \in N^Z} \sum_{a \in A} c_a^{OPk} Q_{nmap} \quad \forall p \tag{4.13}$$

Total transport related costs TC^T for each process realisation p are calculated as a function of (i) the distance traveled ρ_{nm} between upstream nodes n and downstream nodes $m \in M^N$, (ii) the number of batches Q at scale a and process realisation p shipped from n to $m \in M^N$, (iii) a cost parameter for transport c^T.

$$TC_p^T = c^T \sum_{n \in N^M} \sum_{m \in N^M} \sum_{a \in A} \sum_{i \in I} \rho_{nm} Q_{nmap} \quad \forall p \tag{4.14}$$

4.2.3.5 Logic constraints

Logic constraints govern the construction of feasible supply chain network structures. First, a link between an upstream node n and downstream node $m \in M^N$ is established only if the upstream node n exists.

$$X_{nm} \leq Y_n \quad \forall n \in N, m \in M^N \tag{4.15}$$

Where X is the binary variable that marks the establishment of a link between nodes n to $\in M^N$ and Y is the binary variable that marks the existence of a node in location $n \in N$.

Sourcing constraints ensure that each node is sourced by at least one route via the binary variables X. In the case of storage nodes $m \in N^K$, the node exists if the binary variable $Y = 1$ (Eq. 4.16). In the case of demand zones $m \in N^Z$ the nodes are sourced at all time, with the R.H.S. of Eq.5.15 equal to 1.

$$\sum_{n \in N^J} X_{nm} \geq Y_m \quad \forall m \in N^K \tag{4.16}$$

$$\sum_{n \in N^K} X_{nm} \geq 1 \quad \forall m \in N^Z \tag{4.17}$$

In this problem, we consider that at most one scale per plant $n \in N^J$ can be installed. Hence, the following constraint enforces that a scale a is selected only if the manufacturing node exists. Furthermore it enforces the selection of only 1 scale of manufacturing, since the L.H.S. summation cannot take values larger than 1:

$$\sum_{a \in A} V_{na} \leq Y_n \quad \forall n \in N^J \tag{4.18}$$

Where V is the binary variable that marks the selection of scale a in location $n \in N^J$.

Similarly, a production line l in process section s at scale a is established only if the corresponding scale is established at that location $n \in N^J$ and time t:

$$Z_{nasl} \leq V_{na} \quad \forall n \in N^J, a, s, l \tag{4.19}$$

Where Z takes a value of 1 if a line l is established in process section s and at scale a. Product specific equipment installation can occur only if the manufacturing line is available and installed via the binary variable Z. Therefore:

$$M_{nasl} \leq Z_{tnasl} \quad \forall n \in N^J, a, s, l \tag{4.20}$$

Where M marks the product-specific equipment installation in node $n \in N^J$, at scale a, process section s and line l.

In a similar fashion, product allocation W on the L.H.S. of Eq. 4.21 can take value of 1 only if the production line is suitable for production of the product,

which is modelled via $M = 1$. The binary variable W is scenario p dependent and is considered a *second-stage* decision.

$$W_{naslp} \leq M_{nasl} \quad \forall n \in N^J, a, s, l, p \tag{4.21}$$

4.2.3.6 Transport constraints

Transport constraints regulate the maximum and minimum achievable flows between upstream nodes n and downstream nodes $m \in M^N$ in the supply chain network.

$$Q^{MIN} X_{nm} \leq Q_{nmaip} \leq Q^{MAX} X_{nm} \qquad \forall n \in N, m \in M^N \tag{4.22}$$

Where Q^{MIN} and Q^{MAX} correspond to lower and upper bound parameters for the flow variables Q if the link is established via $X = 1$. Instead, these are enforced to be equal to 0 if $X = 0$.

4.2.3.7 Manufacturing constraints

With a similar rationale to transport constraints, time constraints bound the time allocated to production in each process section s, manufacturing line l, process plant $n \in N^J$ in each scenario p.

$$T^{MIN} W_{naslp} \leq T_{naslp} \leq T^{MAX} W_{naslp} \quad \forall n \in N^J, a, l, s, p \tag{4.23}$$

Where the bounds are activated if the corresponding product allocation variable $W = 1$, otherwise the variable T is forced to be equal to 0.

The number of batches B processed in each process section s, process line l at scale a in location $n \in N^J$ and scenario p are calculated as a function of (i) the time allocation variable T, (ii) process parameter realisations in each scenario p. Specifically, these correspond to the process time taken to collect the first batch in the campaign α in each process section s and the length of the bottlenecks within each process section τ.

$$B_{naslp} = W_{naslp} + 1/\tau_{sp}(T_{naslp} - \alpha_{sp} W_{naslp}) \quad \forall n \in N^J, a, l, s, p \tag{4.24}$$

The above expression (Eq. 4.24) is written to compute the number of batches B in each campaign as a sum of the first batch ($W = 1$) and the remaining $B-1$ batches which are obtained within the time slot $T - \alpha$.

4.2.3.8 Network constraints

The overall material balance can be written for every node in the supply chain network can be written as follows.

$$I_{nap} = Q^{IN}_{nap} - Q^{OUT}_{nap} + P_{nap} - S_{nap} \quad \forall n \in N, a, p \tag{4.25}$$

Where I is the inventory of batches (i.e. accumulation) at scale a for product i held, in node n in scenario p, Q^{IN} and Q^{OUT} denote the inflow of batches from upstream nodes and outflow to downstream nodes, P denotes the amount of material produced and S denotes the amount of material consumed (i.e. sold). This general form of material balance can be re-written for each node type. Effectively, in this single-period formulation, inventory is considered zero ($I = 0$) for all nodes, as the network design problem does not consider accumulation over time of material to prepare for future time periods.

In the case of manufacturing facilities $n \in N^J$, only production and outbound flow occur, whilst transformation of materials is not considered in this model. Therefore, the material balance can be written as:

$$0 = P_{nap} - Q^{OUT}_{nap} \quad \forall n \in N^J, a, p \tag{4.26}$$

Where the production term P corresponds of a summation of the batches processed in the F&F process section through lines l, at scale a, in plant $n \in N^J$ at time t and in scenario p.

$$P_{nap} = \sum_{l \in L^S} B_{naslp|s=F} \quad \forall n \in N^J, a, p \tag{4.27}$$

The outbound flow for manufacturing corresponds to a summation of Q from nodes $n \in N^J$ to storage nodes $n \in N^K$.

$$Q^{OUT}_{nap} = \sum_{m \in M^N} Q_{nmap} \quad \forall n \in N^J, a, p \tag{4.28}$$

Furthermore within a manufacturing plant a set of additional mass balances is written to regulate the flow of batches between process sections s. Hence:

$$0 = \sum_{l \in L^S} B_{naslp} - \sum_{l \in L^S} B_{na,s+1,lp} \quad \forall n \in N^J, a, s, p \tag{4.29}$$

Where J corresponds to the inventory of batches held between process section s upstream and its downstream process section $s+1$. This holds for upstream to downstream process sections and downstream to F&F parts of the process.

In the case of storage facilities $n \in N^K$, material is not produced or consumed and the expression in Eq. 4.25 simplifies to a balance between inbound and outbound flows.

$$0 = Q^{IN}_{nap} - Q^{OUT}_{nap} \quad \forall n \in N^K, a, p \tag{4.30}$$

Where the inbound flows Q^{IN} of storage correspond to a summation of Q from nodes $n \in N^J$ to the storage node $n \in N^K$:

$$Q^{IN}_{map} = \sum_{n \in M^N} Q_{nmap} \quad \forall m \in N^K, a, p \tag{4.31}$$

Conversely, the outbound flows Q^{OUT} of storage correspond to a summation of Q the node $n \in N^K$ to demand zones $n \in N^Z$.

$$Q^{OUT}_{nap} = \sum_{m \in M^N} Q_{nmap} \quad \forall n \in N^K, a, p \tag{4.32}$$

Finally, in the case of demand zones $n \in N^Z$, the overall material balance in Eq. 4.25 simplified to a balance between inbound flows and product sales as follows:

$$0 = Q^{IN}_{nap} - S_{nap} \quad \forall n \in N^Z, a, p \tag{4.33}$$

Where Q^{IN} is computed a summation of Q from upstream storage nodes $m \in N^K$ to each demand zone $n \in N^Z$:

$$Q^{IN}_{map} = \sum_{n \in N^M} Q_{nmap} \quad \forall m \in N^Z, a, p \tag{4.34}$$

The amount of doses of LV product supplied to each demand zone corresponds to the amount of batches S at demand zone $n \in N^Z$ at time period t produced with scale a multiplied by the batch size parameter χ. This quantity in doses

has to be greater than or equal that the demand D on the R.H.S at the demand zone $n \in N^Z$.

$$\sum_{a \in A} S_{nap} \chi_{ap} \geq D_n \quad \forall n \in N^Z, p \tag{4.35}$$

In this study, the demand D is calculated by the following equation based on the total demand for each product D^T:

$$D_n = D^T / \text{card}(N^Z) \quad \forall n \in N^Z \tag{4.36}$$

4.2.3.9 Shortage terms

The sales-demand constraint in the deterministic and stochastic optimisation steps can be modified to account for the possibility of the sales not meeting the target demand for certain uncertainty realisations. This can help avoid infeasible solutions, when the capacity available is not sufficient to cater for the target demand. The shortage variable G is introduced to account for potential shortages at each demand zone $n \in N^Z$ and scenario p.

$$\sum_{a \in A} S_{nap} \chi_{ap} + G_{np} \geq D_n \quad \forall n \in N^Z, p \tag{4.37}$$

This requires a modification in the objective function with the introduction of a penalty term for supply shortage.

$$z = \sum_{p \in P} \phi_p (TC_p^{CAPj} + TC_p^{OPj} + TC_p^{OPk} + TC_p^T + \psi \sum_{n \in N^Z} G_{np}) \tag{4.38}$$

Through this reformulation, the objective of the problem is to minimise costs and backlog where possible. The ψ parameter is selected to be large enough to promote maximisation of supply and avoid solutions where it results less costly to create shortages.

4.2.4 Stochastic simulation framework

The performance of each optimisation solution can be tested via stochastic simulation. The optimisation is solved once, whether deterministically or stochastically with $P^O = \{P_1, P_2, ...P_O\}$, to obtain the scenario-independent design decisions. These are fixed and the solution is simulated over a number of runs.

A fixed number of samples $K = 2^6$ is taken to approximate the probability distribution of the uncertain input variables and compute mean output values (Figure 4.1c). An equivalent sampling procedure as presented for the random sampling in Section 5.2.2 is implemented over a set of scenarios $P^S = \{P_1, P_2, ...P_K\}$. It is worth noting that in this performance assessment step, we are seeking to test the *out-of-sample* stability of optimisation solutions. Hence, the integers are within the range 0 to Ω, which allows sampling elements within the full matrix Θ of the original datasets generated via UA and GSA in Chapter 3.

$$X_{\text{test}} \in \{1, 2, \ldots, \Omega\}^{K \times d} \tag{4.39}$$

$$\xi_{kd} = f(X_{\text{test}}, \Theta, \Gamma_\xi) \quad \forall k \in P^S, d \tag{4.40}$$

In each run, the key output metric corresponds to the service level S, which quantifies the capacity of each network structure in meeting target demands. For each Monte Carlo run r, the proportion of service level samples that meet the target demand is recorded, σ_r.

This process is repeated over a number of runs $R = 100$ to obtain values of the expected service level σ_r. At each run the standard error (S.E.) of the mean and overall mean δ is computed and updated. After termination, if S.E.$\leq 5\%$ of the overall mean δ, the output is considered to have converged. Effectively, the number R is an arbitrary number which could be varied alongside K, as long as the convergence criteria is satisfied. The tuning of these two parameters to speed up computational times is not explored in this Chapter.

Ultimately, the performance metric δ corresponds to a simulated service level under uncertainty of each candidate network (i.e. a percentage demand satisfaction). Hence, the closer δ is to 100%, the better the performance of the optimisation solution under uncertainty. The algorithm implementation is summarised in Algorithm 1.

4.2.5 Computational resources

The framework is built in Python 3.11 implementing the PYOMO package. The MILP is solved with GUROBI v11 on a Intel(R) Core(TM) i7-10700 CPU @ 2.90GHz machine with 64.0 GB RAM. Model statistics are displayed in Table 4.1. It is worth noting that the displayed optimisation computational times

Algorithm 1 Monte Carlo Algorithm

Require: Number of samples O in d-dimensions ($d = A \times I$)

Obtain scenario-dependent input parameters ξ_p
Obtain scenario-independent input parameters β
Solve optimisation problem
Record first-stage decisions x
Record second-stage decisions y_p
for $r \leftarrow 1$ to $R = 100$ **do**
    $k \leftarrow 0$
    for $k \leftarrow 1$ to K **do**
        Fix design decisions x
        Simulate design for parameters ξ_k
        Record second-stage decisions y_k
        Compute output level σ_r
    end for
end for
Compute average of output levels σ_r, δ
Compute S.T.D. and S.E.
if S.E. $\leq 5\%$ **then**
    Convergence $\leftarrow$ True
else
    Convergence $\leftarrow$ False
end if

Table 4.1: Computational statistics of the optimisation

Model	**CPU**	**Constraints**	**Integer (Binary)**	**Continuous**	**Gap**
Deterministic	$<$ 1s	1640	806 (661)	519	0%
Stochastic ($O = 2^2$)	1s	5710	2321 (1741)	2051	0%
Stochastic ($O = 2^3$)	2.5s	11136	4341 (3181)	4095	0%
Stochastic ($O = 2^4$)	7.5s	21988	8381 (6061)	8184	0%
Stochastic ($O = 2^5$)	43s	43692	16461 (11821)	16359	0%
Stochastic ($O = 2^6$)	60s	87100	32621 (23341)	32711	0%
Stochastic ($O = 2^7$)	172.5s	173916	64941 (46381)	65416	0%
Simulation run ($K = 2^6$)	$1s \times 2^6$	1640	806 (661)	519	0%

may vary according to the values selected within the sample set P^O. A single simulation run requires CPUs of 1s$\times K$, which in this case corresponds to $K = 2^6$. In this study, the number of runs is set to 100 and convergence with respect to the standard error (S.E.$\leq$ 5%) is checked after termination of the algorithm. Effectively, each run is independent and this could allow a parallelisation of the procedure thus reducing end-to-end computational times by $\times$100 and allowing for larger K values.

4.2.6 Case study description

The case study considers manufacturers producing lentiviral vectors for gene therapy applications (LV). This serves a case to illustrate planning for a single-product manufacturing scenario. The input parameters for each case study are presented in Appendix C. In this study, we consider a range of demands D^T:

1. Phase I clinical demand: 5,000 doses per year
2. Phase II/III clinical demand: 10,000 doses per year
3. Commercial demand: 20,000 doses per year

4.3 Results & discussion

4.3.1 Optimisation results

The presented framework can help assess optimisation strategies under uncertainty and identify those most suited for the problem at hand. We compare

deterministic optimisation cases and a stochastic optimisation approach.

In deterministic optimisations across demands and process realisations, the optimisation yields supply chain networks that meet the target demand (Figure 4.3a). Information on manufacturing capacity of the networks is displayed in Figure 4.4 and Table 4.2. As demand increases, the process volumetric scale increases due to the effect of economies of scale as well as number of lines, with a priority of scaling-up rather than scaling out. This translates into larger capital investments and process costs (Figure 4.3b-c).

More batches are scheduled and subsequently more time is allocated to production (Figure 4.3d-e). This is not the case for the best-case scenario realisation, where a larger capacity actually requires fewer batches to be produced. The average production line utilisation is not strictly increasing (Figure 4.3f). For a median or worst case scenario or stochastic realisation, switching from a medium to high demand results in the installation of more lines (Table 4.2). The average line utilisation is computed as total USP time divided by the installed lines, therefore an increasing number of lines lowers or maintains the average utilisation.

In the best-case scenario, a 200L facility is sufficient to satisfy a low demand (Figure 4.4a) and scale up to medium and higher demand requires 1,000L facilities. The batches required to meet these are the lowest due to the assumption that batch size χ is assumed at its upper bound. Furthermore the time required to meet the target demands is mostly lower (0-250 days) due to the assumption that each batch manufactured requires batch times τ and α at their minimum value bound. Similarly, the cost of each manufacturing parameter for variable c^{VARjl}, labour c^{LABjl} and facility-related c^{FACjl} costs at each scale is assumed to be at its best-value, which is the lower bound. The combination of lower cost coefficient and the lower annual throughput results in lower costs overall.

For each target demand, assuming process capacity parameters at the the median and worst case scenarios, results in a larger number of batches (Figure 4.3d). This is a result of a less optimistic batch size χ. Similarly, the times estimated for production increases rapidly above 350 days due to less optimistic process times α and longer bottlenecks τ (Figure 4.3e). Costs and capital investments also increase as larger capacity is installed and cost coefficients are assumed to be higher. In particular, the assumption of worst-case process realisation results in a solution which maximises utilisation of the USP (Figure 4.4i), with 4 lines at

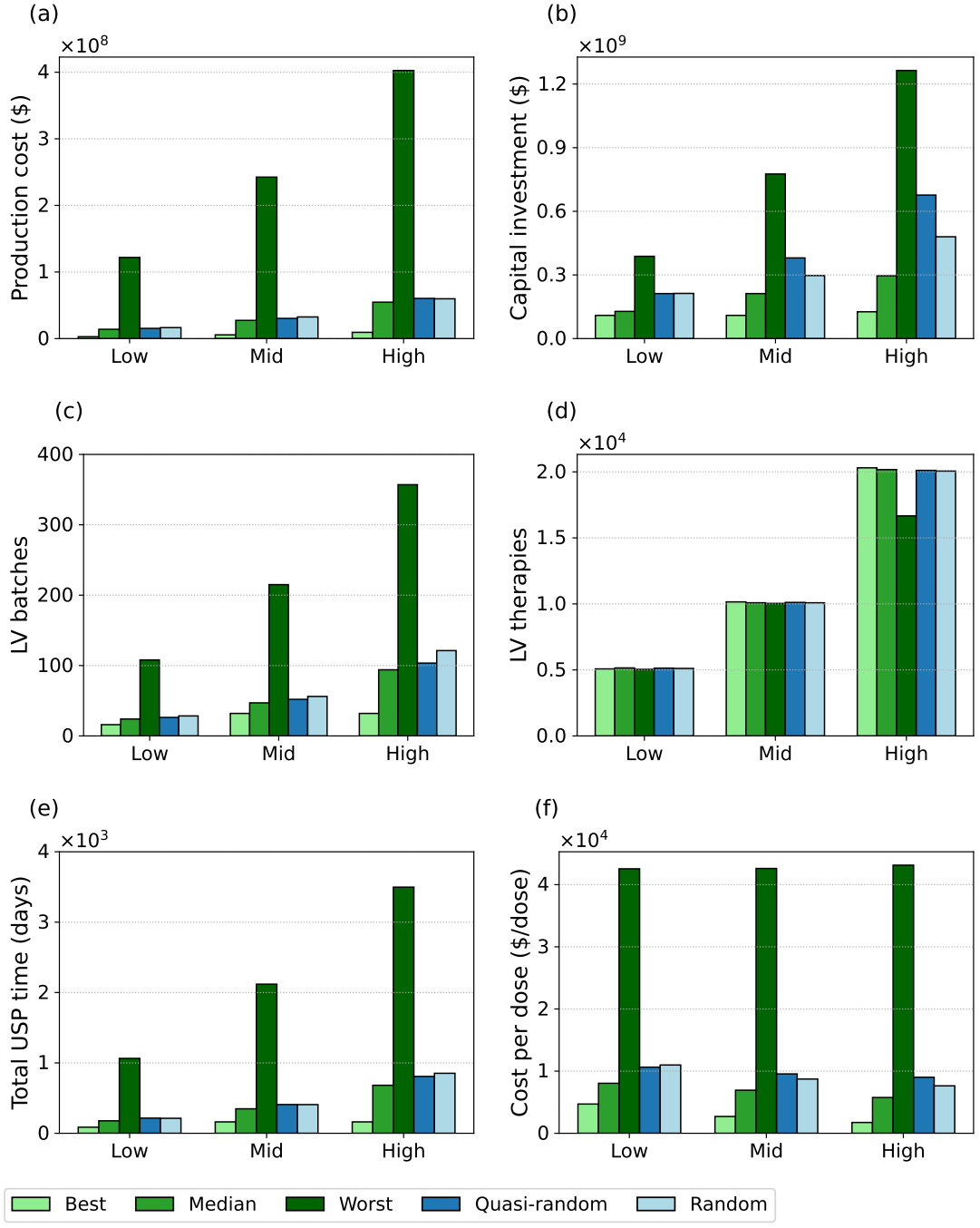


Figure 4.3: Network KPIs for single-product case: 3 different demand realisations (low, medium, high) and 4 different process realisations (best-case, median case, worst-case, stochastic-case) for LV (a-f)

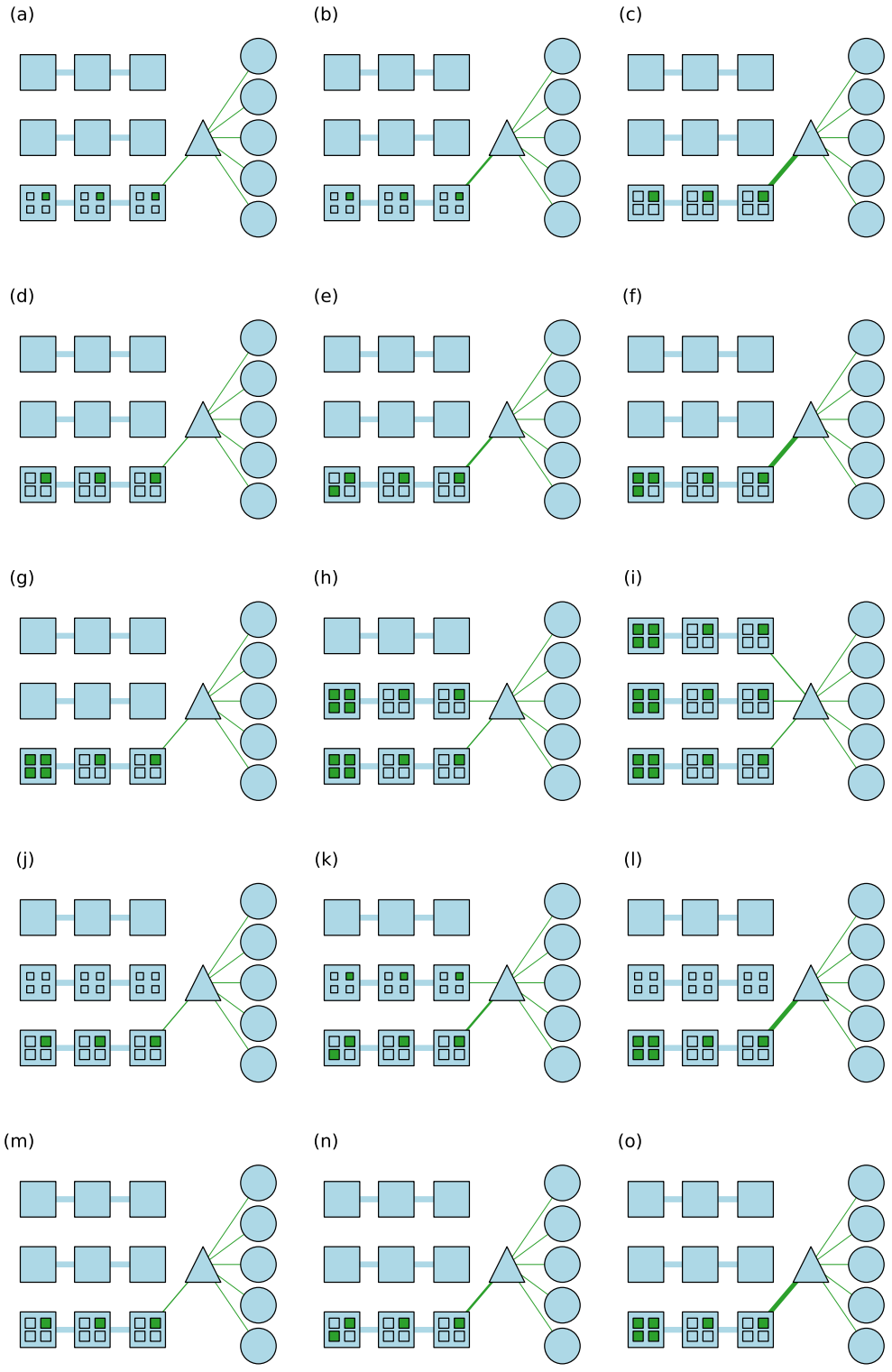


Figure 4.4: Supply chain structures and product allocation for different optimisation strategies and demands. *Best* case (a) low, (b) mid, (c) high demand, *median* case (d) low, (e) mid, (f) high demand, *worst* case (g) low, (h) mid, (i) high demand, *quasirandom* case (j) low, (k) mid, (l) high demand, *random* case (m) low, (n) mid, (o) high demand. Product allocated (■), inactive lines (□).

Table 4.2: Summary of network structures and scales for single-product deterministic case. 3 different demand realisations (low, medium, high) and 3 different process realisations (best-case, median case, worst-case) for LV.

	Demand D^T								
	Low			*Medium*			*High*		
	Node	Scale	Lines	Node	Scale	Lines	Node	Scale selected	Lines
Best	*j1*	200	1	*j1*	200	1	*j1*	2000	1
	j2	-	-	*j2*	-	-	*j2*	-	-
	j3	-	-	*j3*	-	-	*j3*	-	-
Median	*j1*	2000	1	*j1*	2000	2	*j1*	2000	3
	j2	-	-	*j2*	-	-	*j2*	-	-
	j3	-	-	*j3*	-	-	*j3*	-	-
Worst	*j1*	2000	4	*j1*	2000	4	*j1*	2000	4
	j2	-	-	*j2*	2000	4	*j2*	2000	4
	j3	-	-	*j3*	2000	3	*j3*	2000	4
Stochastic (quasi-random)	*j1*	2000	2	*j1*	2000	4	*j1*	2000	4
	j2	-	-	*j2*	-	-	*j2*	1000	1
	j3	-	-	*j3*	-	-	*j3*	-	-
Stochastic (random)	*j1*	2000	2	*j1*	2000	2	*j1*	2000	4
	j2	-	-	*j2*	-	-	*j2*	-	-
	j3	-	-	*j3*	-	-	*j3*	-	-

2,000L scale in each active manufacturing node (Table 4.2), whilst not meeting the demand and creating therapy backlogs (Figure 4.3a). Additionally, worst-case deterministic optimisation results in 10-fold larger costs compared to the stochastic counterparts (Figure 4.3b-c).

In the stochastic cases, the optimisation finds a solution which is not as conservative as the worst case scenario, which highlights that the P^O samples generated for process scenarios cluster more around the central part of the distribution, possibly near the median value. We find that the network structure differs if the stochastic optimisation is based on random or quasi-random sampling. In the random sampling-based case (Figure 4.4n-o) for medium and high demands production is allocated to smaller facilities which result in lower costs (Figure 4.3b-c). In the quasi-random case, costs result higher than both the

random case and the median case, and this would be due to better coverage of the uncertainty space towards worst-case values.

It is worth noting that the stochastic network designs are found to be influenced by the samples drawn. This is an indication that we can obtain different *iso-cost* designs of comparable performance and leveraging on the same number of scenarios. In Figure 4.4, we report example structures which are simulated in Section 4.3.2. However, it is crucial to study the variability of network design to the scenarios selected.

4.3.2 Simulation results

As found in Section 4.3.1, over-designing facilities through a worst-case optimisation approach results in larger costs than those for the deterministic solutions at the median level or the stochastic cases. In this context, supply chain simulation is used as a tool to assess the actual benefit of adopting a worst-case deterministic or stochastic optimisation approach to mitigate shortage risks versus other uncertainty integration strategies.

Whilst network structures obtained with stochastic optimisation may vary according to the samples drawn and achieve equivalent objective function values, the simulation remains a valuable tool to test whether given a design, this achieves the required production targets.

The implementation of the simulation aims to construct statistical bounds on the supply chain network performance. In doing so at every Monte Carlo run, the framework is computing the cumulative distribution as shown in Figures 4.5a-c and recording the value of σ_r, which corresponds to the proportion of samples below the target of 100% demand satisfaction. At the same time, we report the cumulative distributions of costs in Figure 4.5d-f. This allows a quantification of the cost-effectiveness of each solution under uncertainty. For the service level KPI, a good solution minimises the probability that the service level is below the target. Conversely, for the cost related KPI, a good solution maximises the probability of the cost per dose being below the target value.

For all demand scales considered, the best-case scenario structures perform the worst under uncertainty, with percentage probabilities of not meeting target demands larger than 50%, 96%, 90% respectively (Figure 4.5a-c) and the smallest probabilities of being below the cost per dose target (Figure 4.5d-f). The median

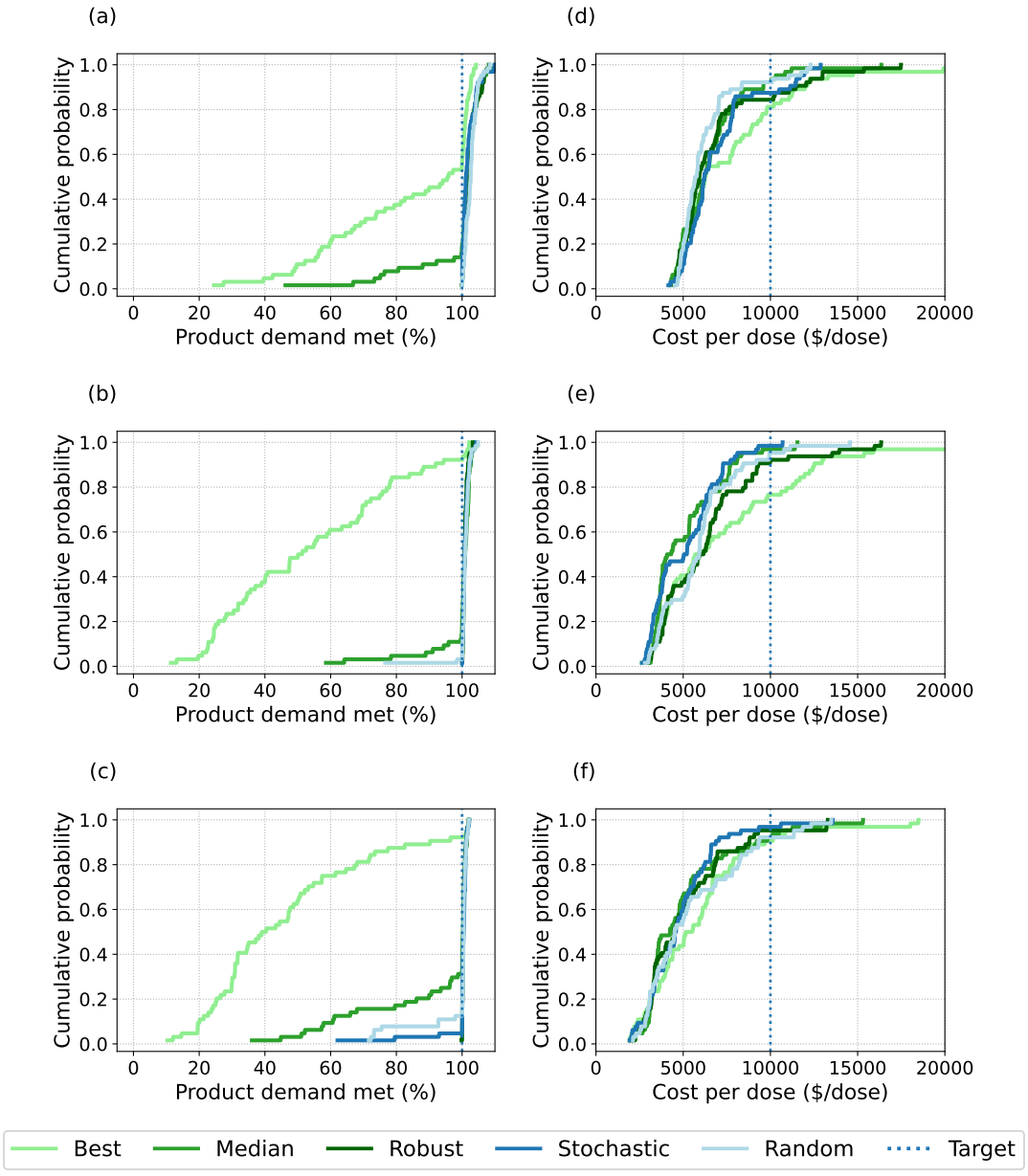


Figure 4.5: Simulation via random sampling ($P^S = 2^6$ samples) for single-product network structures generated for (a) low, (b) mid, (c) high demand for LV and cost per dose for (d) low, (e) mid, (f) high demand.

case performs under uncertainty with probabilities of unmet demands within the range of $10 - 30\%$. Conversely, optimising for the worst case scenario ensures a probability of 1 to meet the demand for all process realisations. It is found that in low and medium demand cases, the cost probability is lower for the best case solution (Figure 4.5d-e), highlighting that the cost-effectiveness for this solution is probabilistically better within this single iteration, however at the expenses of supplying below the target demand.

The stochastic cases are found to perform better than the deterministic solution at the best and median case with probabilities of 0 of supplying below the target demand. Effectively, designs generated via stochastic programming perform similarly to the worst-case designs for all demand cases (Figure 4.5a-c). This highlights that the stochastic formulation yields good solutions that ensure a high probability of meeting target demands whilst resulting in similar simulated costs compared to the worst-case solution (Figure 4.5d-f).

Upon a more thorough comparison between stochastic optimisation based on random or quasi-random sampling, the sampling approaches yield a comparable simulation performance. Specifically, in the single simulation run presented in Figure 4.5a-b quasi-random sampling performs similarly to the random-based counterpart, with a probability of not meeting the target demand of approximately 0%. Instead, in Figure 4.5c, random sampling has a slightly larger probability of approximately 10% of supply below the target demand. This could be an instance of quasi-random sampling achieving a better coverage of the uncertainty space for the same number of samples.

Multiple runs of the Monte Carlo simulation help quantify a mean output level δ for each optimisation strategy with a parallel sampling approach. In this step, we report results of the simulated service level and cost (Figure 4.6a-b), As can be seen in Figure 4.6a, the results of the Monte Carlo algorithm highlight a consistent worse performance under uncertainty for the best-case deterministic optimisation, followed by the median-case optimisation. As expected, worst-case optimisation always results in $\delta \approx 100\%$. Similarly, stochastic optimisation approaches also achieve a mean service level which is consistently higher than the median and best-case deterministic optimisations $\delta \geq 84\%$. In terms of costs, the best-case solutions lead to consistently to lower probabilities of a cost per dose below the target value. Instead, other solutions result in consistently larger probabilities (80-100 %) of being below target and better cost effectiveness.

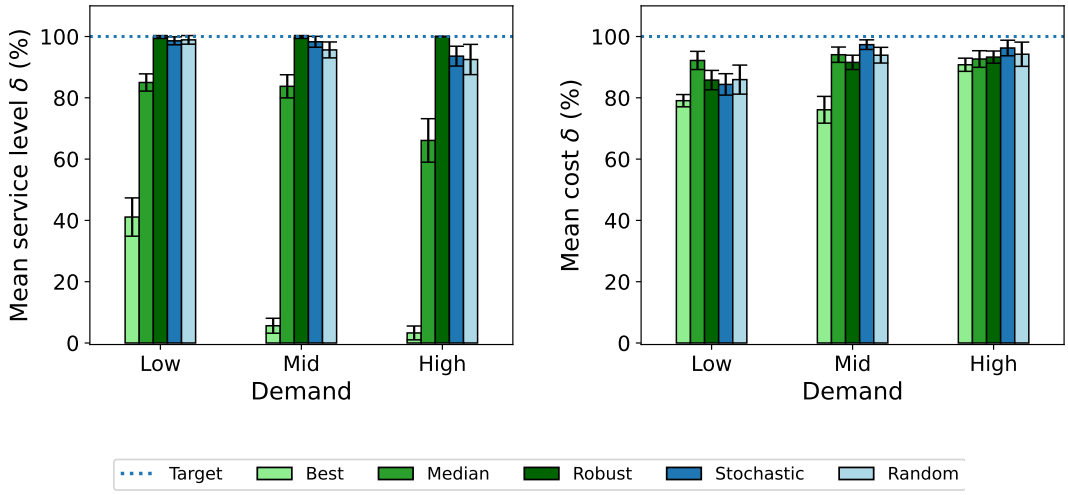


Figure 4.6: Multiple-runs simulation for different designs obtained with optimisation strategies: (a) mean probability of meeting target demand, (b) mean probability of cost below target

Effectively, the simulation results highlight a trade-off between integrating process uncertainty and the achievable service levels, which do not directly correlated to cost metrics. In fact, the presented formulation considers a shortage cost, which is quantified via a penalty term. Refining this to account for an industry-informed cost of missed delivery, a trade-off between the effort of quantifying uncertainty and generating *uncertainty-aware* solutions, and total supply chain costs would emerge. This would be used to inform decision-makers regarding the comprehensive cost- and supply-related risks tied to disregarding process uncertainty integration and motivate the adoption of advanced planning tools.

4.3.3 Scenario tuning

As argued in Section 4.3.1 and 4.3.2, sampling-based stochastic optimisation yields solutions which can achieve reduced costs and perform well under uncertainty. The closed loop simulation can help quantify statistical bounds on the actual performance of different sampling strategies.

We present results as a function of scenarios used in the stochastic optimisation, with the aim of identifying the minimum number of samples, that achieves good values of δ, is expected to minimise computational cost and thus allows the benefits of uncertainty integration to be accurately determined.

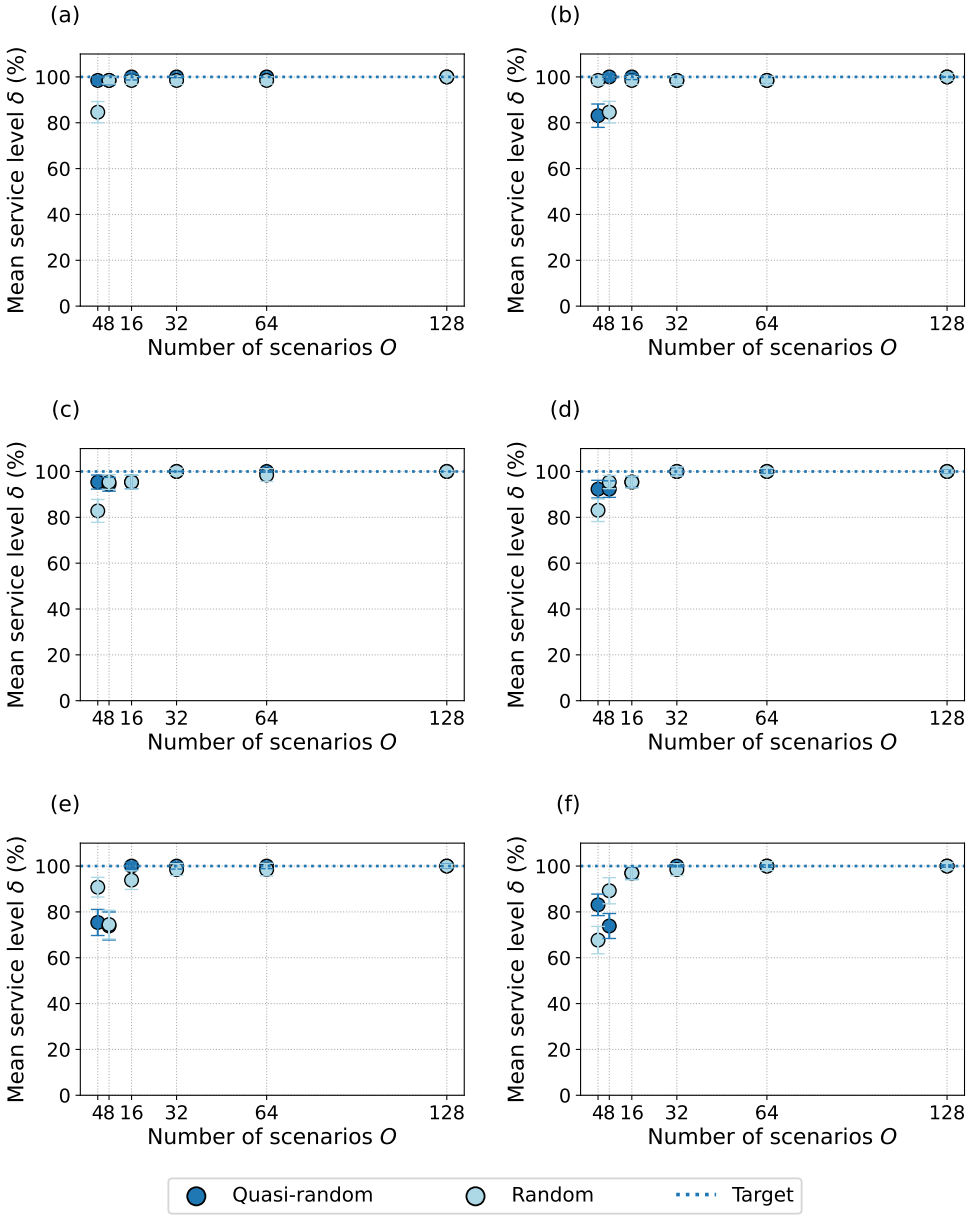


Figure 4.7: Scenario tuning results for Trial 1 and demand (a) low, (c) mid, (e) high and Trial 2 and demand (b) low, (d) mid, (f) high.

We consider scenarios $O = \{4, 8, 16, 32, 64, 128\}$, as quasi-random sampling is based on Sobol sequences, which achieve better coverage if generated as powers of 2. In Figure 4.7a-f, we compare the performance of the scenario generation-strategy for different demand scales (low, medium, high) and repeating the exercise for two independent trials, Trial 1 and Trial 2. Given 4 quasi-random scenarios sampled for the stochastic program, the average percentage probability of meeting the demands ranges within $79 \pm 2\%$ (Figure 4.7e) and $99 \pm 0.2\%$

(Figure 4.7a). Hence, in some demand and trial instances, this results lower than the random counterpart. It is also expected that at low number of scenarios, the output structure selected by the optimisation differs based on the samples scenarios. Given 8 quasi-random scenarios, the performance of the solutions upon simulation remains variable within the range of $73 \pm 2\%$ (Figure 4.7e-f) and $99 \pm 0.1\%$ (Figure 4.7b). This highlights that the randomisation either through quasi-random samples shuffling or in the actual randomness of samples in the random case may hinder the coverage of the uncertainty space through a limited set of scenarios. In addition, this is an indication of a possible variability in the underlying network structure obtained which is then simulated.

It is found that as the number of samples increases both sampling strategies tend to converge to good δ values. As O increases to 16 and above, both random and quasi-random samples perform near equivalently in meeting 100% of the demands upon simulation. Hence, one can identify a minimum number of randomly or quasi-randomly generated scenarios that should be included in the stochastic program to achieve a good accuracy and benefit. This exercise can be repeated in cases considering a larger number of independent processes and would help identify the number of $n-$dimensional samples required to achieve good coverage of the uncertainty space. In addition to this, it is crucial to investigate whether the underlying network structures is susceptible to the set of drawn samples at a high number of scenario.

4.4 Conclusions

In this Chapter, we assess and explore the potential of optimisation under uncertainty strategies in network design problems. We compare worst-case optimisation approaches and stochastic programming with deterministic counterparts that do not account for process uncertainty through a simulation approach.

Firstly, we demonstrate that optimal design obtained through worst-case and stochastic programming outperform deterministic designs, which ignore process uncertainty, with respect to expected service levels and costs. Thus, the benefit of early-stage process uncertainty integration in supply chain design is quantified. The optimisation results highlight that worst-case optimisation leads to overly conservative decisions which result in 10-fold larger costs compared to the stochastic counterpart. Upon simulation, the stochastic network designs per-

form near-equivalently to the worst-case designs, whilst corresponding to lower expected costs, which demonstrates that these perform near equivalently upon implementation, making both options valid approaches for integration.

Secondly, we develop a strategy to pinpoint a minimum number of sampled scenarios which represent the process uncertainty space quantified through Chapter 3. There remains open question related whether the underlying structure is consistently the same; however, results highlight that the identified minimum set of scenarios yields a consistent simulated performance, indicating that the framework provides guarantees of good coverage of the uncertainty space. These findings underpin the framework discussed in Chapter 5, which introduces additional products and leverages on the same stochastic optimisation strategy. We explore flexibility to unknown unknowns under unconstrained and constrained design decisions. In the latter case, our approach can guarantee a good description of the uncertainty space for a fixed structure, thus allowing further conclusions to be drawn.

Chapter 5

Quantification of flexibility under uncertainty

Outputs of this chapter have been disseminated in:

- **Sarkis, M.**, Shah, N. & Papathanasiou, M.M. Resilient multi-product next-generation pharmaceutical supply chains: quantification of flexibility under uncertainty. *Submitted to Journal of Advanced Processing & Manufacturing*. Under review.

5.1 Introduction

Investment planning in the context of emerging pharmaceutical sectors requires rapid scale-up decisions to meet growing demands. In fact, as the sector transitions to flexible, modular multi-product manufacturing facilities, with platforms that should readily adapt to demand changes, integrating new product approvals, scaling up capacity and orchestrating between different product campaigns.

Adapting to demand changes is not a straightforward exercise as one should consider the time-scale of recourse actions available. In Chapter 2, a number of contributions which use stochastic programming in an investment planning context was presented, specifically under demand uncertainty. Gatica et al. (2003a) use stochastic optimisation to obtain investment planning decisions considering clinical trial outcome uncertainty. With a similar rationale, Tsang et al. (2007b) focus on capacity planning for a range of vaccine applications and analyse different risk metrics to consider within the decision-making. Sundaramoorthy et al. (2012b) focus on capacity planning under uncertainty and propose a decomposition method to address the computational complexity that arises from problem instances embedding large scenario trees. Further studies have focused on investment planning under demand uncertainty for large-scale pharmaceutical supply chains, tackling location-allocation and reverse logistics problems through industrially relevant case studies (Amaro and Barbosa-Póvoa, 2008, 2009; Mestre et al., 2015; Zeballos et al., 2014).

Efforts have also developed production planning models tailored to biopharmaceutical manufacturing complexities. Lakhdar et al. (2005), Siganporia et al. (2014) develop mid-term planning formulation, considering additional complexities, including perfusion or fed-batch operations, facility expansion decisions and solve the problem with a rolling horizon approach. Furthermore, Liu et al. (2014) consider the problem of production planning combined with maintenance planning, integrating the complexity of resin consumables time-dependent decay in the decision-making. The problem of production planning is also tackled in Lakhdar et al. (2007) with the integration of process uncertainty. Given time-varying target demands, the proposed formulation integrates uncertainty distributions for fermentation titers via an individual chance constraint (Charnes and Cooper, 1959; Küçükyavuz and Jiang, 2022). In this instance, capacity expansion decisions are not considered and the model focuses on campaign planning. In

doing so, the production plans generated ensure demand satisfaction for different probability levels specified *a priori*.

Effectively, demands for vaccines for unforeseen pandemics can be challenging and highlight the unpreparedness of pharmaceutical networks in accommodating vaccine production alongside established products. Given unknown timings of realisation of demands, the flexibility of manufacturing capacity to cater for vaccine production in the short-term becomes a crucial metric for informed investment planning. However, this flexibility in the context of process platforms under development should integrate information on the underlying process uncertainty in order to reliably estimate product-specific process capabilities.

In Chapter 4, the optimisation considers *single-product* manufacturing networks and focuses on an lentivirus (LV)-based gene therapy case study. This chapter presents a framework for the quantification of flexibility and reliability of manufacturing capacity in catering for *multi-product* demands, from emerging targeted healthcare to vaccine production. In doing so, it becomes crucial to integrate process uncertainty knowledge acquired *a-priori* to enhance the resilience of the decision making, previously quantified in Chapter 3. The framework augments the single-product formulation presented Chapter 4 to multiple products and embeds a two-stage stochastic optimisation step that solves an investment planning problem with a pre-specified probability level for each product. This step relies on an approximation of the uncertainty space, which leverages on the scenario approximations proposed in Chapter 4. Therefore, the accuracy of the approximation is tested via a Monte Carlo method which provides an *a posteriori* insight into the probabilistic performance of each candidate solution.

In this chapter, we use this to quantify the accuracy of a scenario-based chance constraint satisfaction. By introducing chance-constraints in the optimisation, we can quantify cost-reliability trade-offs, where the cost of over-designing facilities and the probability of meeting the demands of different products are the quantities of interest. We select a case study from the viral vector sector, namely LV-based gene therapy and a competing adenoviral (AdV) based vaccine. The framework is transferable to any other emerging therapeutic and vaccine of interest.

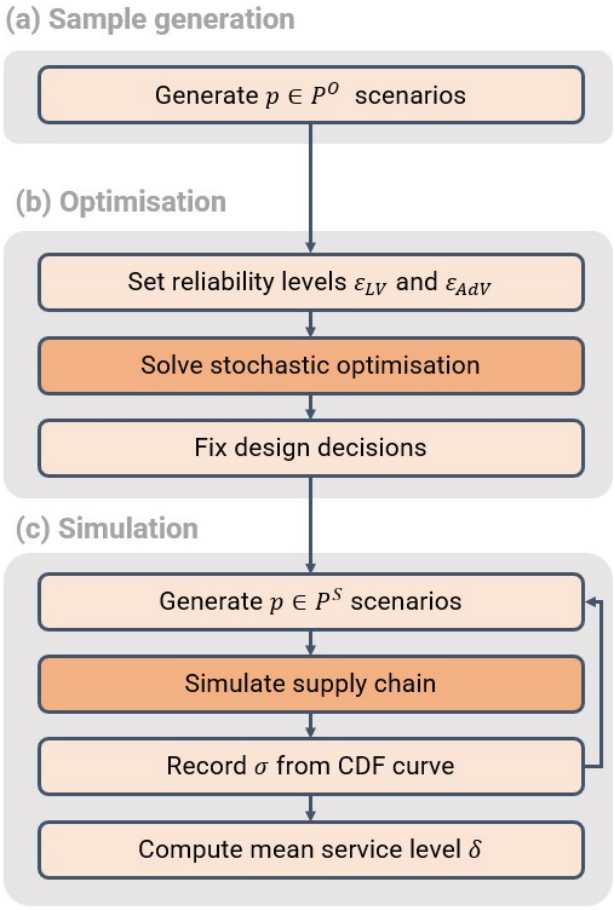


Figure 5.1: Optimisation-simulation framework: (a) process scenarios are generated for the optimisation (b) the optimisation is formulated as a deterministic problem or a stochastic program, (c) the simulation scheme is implemented to quantify the statistical performance of each optimisation solution.

5.2 Methodology

The framework presented consists of (i) a process scenario generation step, (ii) an optimisation step for network design and (iii) a simulation step for statistical performance quantification (Figure 5.1). The following sections illustrate the methodology adopted in the development of the optimisation-simulation framework and discuss the case studies used to illustrate the capabilities of the tool.

5.2.1 Problem statement

The problem considered in this study focuses on capacity planning and is based on the network superstructure illustrated in Figure 5.2. The objective of the

planning problem is to minimise total supply chain capital and operating costs and determine the following decisions:

- a network structure and links between nodes;
- investment in manufacturing and storage facilities;
- scales of manufacturing and number of parallel lines;
- production levels at each manufacturing node;
- transportation flows between existing nodes;
- supplied quantities at each demand node.

To do this, the optimisation considers a range of inputs including:

1. a set I of products i;
2. a set N of nodes n;
3. a set A of manufacturing scales a;
4. a set S of manufacturing process sections s;
5. a set L^S of manufacturing process lines l;
6. a set P of manufacturing process scenarios p;
7. datasets of scale-dependent batch sizes, batch times and costs;
8. data on distances between each node n;
9. data on transportation costs;
10. data on storage costs;
11. demand of product $i \in I$ in demand zones $n \in N^Z$

The underlying difference from the previous Chapter 4 in this single-period capacity planning problemis the introduction of set I of products, comprising LV and AdV. This translates in modified equations accounting for product-specific data and additional constraints to account for change-over and retrofitting.

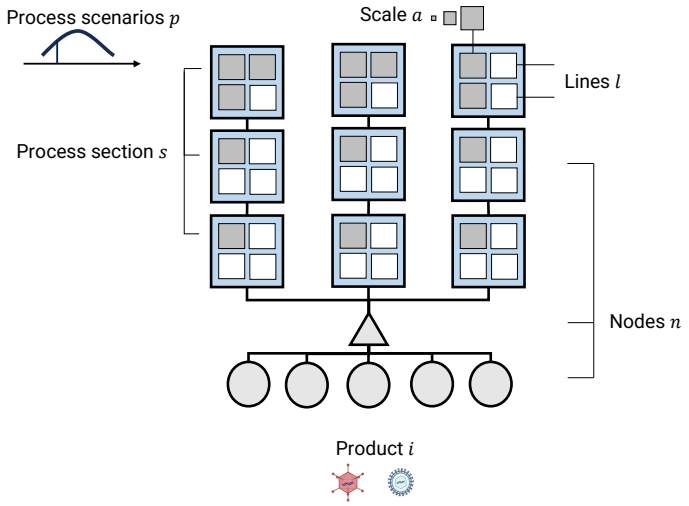


Figure 5.2: Superstructure for optimisation of multi-product i case study. Nodes n comprise manufacturing nodes (square), storage (triangle) and demand zones (circles). Each manufacturing node is modelled as a set of process sections s comprising parallel manufacturing lines l at scale a. Process scenarios p are generated as input datasets for the optimisation-simulation framework.

5.2.2 Process scenario generation

The first step within the framework (Figure 5.1) is the integration of process scenarios as inputs for the planning problem. The aim of this is to assign values to the vector of process inputs ξ, comprising batch sizes χ, process time α, bottleneck length τ and cost parameters c. In this Chapter, the sampling method considers a $d-$dimensional uncertainty distribution of $d = A \times I$, thus considering uncertainty analyses conducted for the two products, LV and AdV within set of products $i \in I$.

Hence, as illustrated in the previous Chapter 4 matrix X_{Sobol} of $O \times d$ integers between 1 and ω. This ensures that the random integers correspond to elements within the submatrix θ.

$$X_{\text{Sobol}} \in \{1, 2, \ldots, \omega\}^{O \times d} \tag{5.1}$$

The matrix of integers X_{Sobol} generated based on low-discrepancy Sobol' sequences, θ are used to assign the respective values for the general ξ parameters based on the corresponding value matrix Γ for the parameter:

$$\xi_{pd} = f(X_{\text{Sobol}}, \theta, \Gamma_\xi) \quad \forall p \in P^O, d \tag{5.2}$$

5.2.3 Optimisation formulation

The optimisation step in Figure 5.1 considers a single-period multi-product planning problem of LV-based gene therapy and AdV-based vaccines. These constitute the set of products $I = \{LV, AdV\}$. The O samples are drawn to generate the uncertain process realisations within $P^O = \{P_1 ... P_N\}$. The sampling procedures implemented are described in Section 5.2.2. In the following Section 5.2.3.1, the MILP optimisation formulation is presented, integrating manufacturing complexities of multi-product manufacturing.

5.2.3.1 Objective function

The objective of the optimisation is to minimise the total expected cost z of the supply chain (Eq.5.3). Total costs include the scenario-dependent capital costs and operating costs of manufacturing, storage costs and transport costs, multiplied by the respective probability ϕ_p.

$$\min \; z = \sum_{p \in P} \phi_p (TC_p^{CAPj} + TC_p^{OPj} + TC_p^{OPk} + TC_p^{T}) \tag{5.3}$$

Where TC^{CAPj} is the total capital manufacturing cost, TC^{OPj} is the total operating manufacturing cost, TC^{OPk} is the total operating storage cost and TC^{T} is the total transport. These are all scenario-dependent quantities as they are functions of scenario-dependent process realisations p for throughput and cost-related quantities.

It is assumed that all scenarios are equi-probable, given that the samples generated are identically distributed samples (IID). Hence the probability ϕ_p is computed as:

$$\phi_p = \frac{1}{O} \quad \forall p \tag{5.4}$$

5.2.3.2 Capital manufacturing costs

Capital costs of manufacturing (Eq.5.5) are computed as a sum of facility installation costs and product-specific equipment investment costs.

$$TC_p^{CAPj} = \sum_{n \in N^J} C_{np}^{EQj} \quad \forall p \tag{5.5}$$

The term within in the summation of Eq. 5.5 is C^{EQj}, which is a scenario-dependent variable accounting for the total cost of equipment investment, in node $n \in N^J$ and process scenario p.

$$C_{np}^{EQj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} \sum_{i \in I} c_{asip}^{EQjl} M_{nalsi} \quad \forall n \in N^J, p \tag{5.6}$$

The equipment installation cost C^{EQj} is computed as a function of (i) *first-stage* binaries M, which indicate whether equipment for the manufacturing of product i at scale a is available and installed in lines l of section s in plant $n \in N^J$, (ii) scenario-dependent capital cost realisations for parameters c^{EQjl}(Eq.5.6):

5.2.3.3 Operating manufacturing costs

Total operating costs TC^{OPj} (Eq. 5.7) of manufacturing correspond to the summation of annual variable costs C^{VARj}, facility related costs C^{FACj} and labour costs C^{LABj} for each manufacturing plant $n \in N^J$, process realisation p.

$$TC_p^{OPj} = \sum_{n \in N^J} C_{np}^{LABj} + C_{np}^{VARj} + C_{np}^{FACj} \quad \forall p \tag{5.7}$$

Labour costs C^{LABj} for each manufacturing plant $n \in N^J$ and scenario p are computed (Eq. 5.8) as the sum of the (i) labour costs c^{LABjl} parameters for each scale a, process section l, product i and process realisation p multiplied by (ii) batches B respectively which vary with respect to plant $n \in N^J$, scale a, process section s, production line l and process realisation p:

$$C_{np}^{LABj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} \sum_{i \in I} c_{asip}^{LABjl} B_{naslip} \quad \forall n \in N^J, p \tag{5.8}$$

Variable costs C^{VARj} for each manufacturing plant $n \in N^J$ and scenario p are computed (Eq. 5.9) as a multiplication of the (i) variable costs parameters c^{VARjl} for each scale a, product i, process section s and process realisation p by (ii) batches B respectively which vary with respect to plant $n \in N^J$, scale a, process section s, production line l, product i and process realisation p:

$$C_{np}^{VARj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} \sum_{i \in I} c_{asip}^{VARjl} B_{naslip} \quad \forall n \in N^J, p \tag{5.9}$$

Facility-dependent costs C^{FACj} are tied to depreciation and equipment maintenance costs. In Eq. 5.10, facility-dependent costs for every manufacturing plant $n \in N^J$ and process realisation p are computed as multiplication of (i) facility costs parameters c^{FACjl} for each scale a, process section l, product i and process realisation p by (ii) equipment installation-related binaries M respectively vary with respect to plant j, scale a, process section s, production line l and product i.

$$C_{np}^{FACj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} \sum_{i \in I} c_{asip}^{FACjl} M_{nalsi} \quad \forall n \in N^J, p \tag{5.10}$$

5.2.3.4 Logistics-related costs

Total operating costs of storage TC^{OPk} are a function of (i) the batches handled in the process Q, which corresponds to the number of batches of scale a of product i shipped from upstream storage node $n \in N^K$ to downstream demand zone $m \in N^Z$ in scenario p, (ii) the cost parameter c^{OPk} for product i at scale a.

$$TC_p^{OPk} = \sum_{n \in N^K} \sum_{m \in N^Z} \sum_{a \in A} \sum_{i \in I} c_{ai}^{OPk} Q_{nmaip} \quad \forall p \tag{5.11}$$

Total transport related costs TC^T for each process realisation p are calculated as a function of (i) the distance traveled ρ_{nm} between upstream nodes n and downstream nodes $m \in M^N$, (ii) the number of batches Q at scale a, product i and process realisation p shipped from n to $m \in M^N$, (iii) a cost parameter for transport c^T.

$$TC_p^T = c^T \sum_{n \in N^M} \sum_{m \in N^M} \sum_{a \in A} \sum_{i \in I} \rho_{nm} Q_{nmaip} \quad \forall p \tag{5.12}$$

5.2.3.5 Logic constraints

Logic constraints govern the construction of feasible supply chain network structures. First, a link between an upstream node n and downstream node $m \in M^N$ is established only if the upstream node n exists at time t.

$$X_{nm} \leq Y_n \quad \forall n \in N, m \in M^N \tag{5.13}$$

Where X is the binary variable that indicates the establishment of a link between nodes n to $\in M^N$ and Y is the binary variable that indicates the existence of a node in location $n \in N$.

Sourcing constraints ensure that each node is sourced by at least one route via the binary variables X. In the case of storage nodes $m \in N^K$, the node exists if the binary variable $Y = 1$ (Eq. 5.14). In the case of demand zones $m \in N^Z$ the nodes are sourced at all time, with the R.H.S. of Eq.5.15 equal to 1.

$$\sum_{n \in N^J} X_{nm} \geq Y_m \quad \forall m \in N^K \tag{5.14}$$

$$\sum_{n \in N^K} X_{nm} \geq 1 \quad \forall m \in N^Z \tag{5.15}$$

In this problem, we consider that at most one scale per plant $n \in N^J$ can be installed. Hence, the following constraint enforces that a scale a is selected only if the manufacturing node exists. Furthermore it enforces the selection of only 1 scale of manufacturing, since the L.H.S. summation cannot take values larger than 1:

$$\sum_{a \in A} V_{na} \leq Y_n \quad \forall n \in N^J \tag{5.16}$$

Where V is the binary variable that indicates the selection of scale a in location $n \in N^J$.

Similarly, a production line l in process section s at scale a is established only if the corresponding scale is established at that location $n \in N^J$:

$$Z_{nasl} \leq V_{na} \quad \forall n \in N^J, a, s, l \tag{5.17}$$

Where Z takes a value of 1 if a line l is established in process section s and at scale a. Product specific equipment installation can occur only if the manufacturing line is available and installed via the binary variable Z. Therefore:

$$M_{nasli} \leq Z_{nasl} \quad \forall n \in N^J, a, s, l, i \tag{5.18}$$

Where M indicates the product-specific equipment installation for product i, in node $n \in N^J$, at scale a, process section s and line l.

In a similar fashion, product allocation W on the L.H.S. of Eq. 5.19 can take value of 1 only if the production line is suitable for production of the product, which is modelled via $M = 1$. The binary variable W for product i is scenario p dependent and is considered a *second-stage* decision.

$$W_{naslip} \leq M_{nasli} \quad \forall n \in N^J, a, s, l, i, p \tag{5.19}$$

If two products are manufactured in the same manufacturing node $n \in N^J$, scale a, process section s and line l, a change over occurs. To express this, the binary variables N are introduced to take a value of 1 if a change-over has occurred or 0 if otherwise. This is modelled as follows:

$$N_{naslp} \leq \sum_{i \in I} W_{naslip} - 1 \quad \forall n \in N^J, a, s, l, p \tag{5.20}$$

Where the L.H.S takes a value of 1 only if the first term on the R.H,S, is greater than 1.

Furthermore, it is worth considering that in this multi-product formulation, change overs occur throughout the time period t and the sequence of production is not considered. Therefore, the following equations are introduced to account for change-overs occurring only once, assuming that one product (LV) is always produced before the other (AdV).

$$C_{naslip|i=LV} \leq N_{naslp} \quad \forall n \in N^J, a, s, l, p \tag{5.21}$$

$$C_{naslip|i=AdV} \leq N_{naslp} - 1 \quad \forall n \in N^J, a, s, l, p \tag{5.22}$$

5.2.3.6 Transport constraints

Transport constraints regulate the maximum and minimum achievable flows between upstream nodes n and downstream nodes $m \in M^N$ in the supply chain network.

$$Q^{MIN} X_{nm} \leq \sum_{i \in I} Q_{nmaip} \leq Q^{MAX} X_{nm} \qquad \forall n \in N, m \in M^N \tag{5.23}$$

Where Q^{MIN} and Q^{MAX} correspond to lower and upper bound parameters for the flow variables Q if the link is established via $X = 1$. Instead, these are enforced to be equal to 0 if $X = 0$.

5.2.3.7 Manufacturing constraints

With a similar rationale to transport constraints, time constraints bound the time allocated to production of product i in each process section s, manufacturing line l, process plant $n \in N^J$ in each scenario p.

$$T^{MIN} W_{naslip} \leq T_{naslip} \leq T^{MAX} W_{naslip} \quad \forall n \in N^J, a, l, s, i, p \tag{5.24}$$

Where the bounds are activated if the corresponding product allocation variable $W = 1$, otherwise the variable T is forced to be equal to 0. Additional KPIs of manufacturing are the total utilisation of manufacturing capacity, which is computed as a summation of the time T allocated to production of each product i (Eq. 5.25). An overall upper bound can also be imposed via Eq. 5.26.

$$U_{naslp} = \sum_{i} T_{naslip} \quad \forall n \in N^J, a, s, l, p \tag{5.25}$$

$$U_{naslp} \leq U^{MAX} \quad \forall n \in N^J, a, s, l, p \tag{5.26}$$

The number of batches B of product i processed in each process section s, process line l at scale a in location $n \in N^J$ and scenario p are calculated as a function of (i) the time allocation variable T, (ii) process parameter realisations in each scenario p. Specifically, these correspond to the process time taken to collect the first batch in the campaign α in each process section s, the length of the bottlenecks within each process section τ and changeover downtime due to multi-product manufacturing.

$$B_{naslip} = W_{naslip} + 1/\tau_{sp}(T_{naslip} - \alpha_{sp} W_{nasip} - \kappa C_{tnasip}) \quad \forall n \in N^J, a, s, i, p \tag{5.27}$$

The above expression (Eq. 5.27) computes the number of batches B in each campaign as a sum of the first batch ($W = 1$) and the remaining $B - 1$ batches which are obtained within the time slot $T - \alpha - \kappa$. The latter term ensures that downtime due to change-over is accounted for in the case where $C = 1$.

5.2.3.8 Network constraints

The overall material balance for every node in the supply chain network can be written as follows:

$$I_{naip} = Q^{IN}_{naip} - Q^{OUT}_{naip} + P_{naip} - S_{naip} \quad \forall n \in N, a, i, p \tag{5.28}$$

Where I is the inventory (i.e. accumulation) of batches at scale a for product i held in node n in scenario p, Q^{IN} and Q^{OUT} denote the inflow of batches from upstream nodes and outflow to downstream nodes, P denotes the amount of material produced and S denotes the amount of material consumed (i.e. sold). This general form of material balance can be re-written for each node type.

First, in the single-period problem there is no accumulation allowed, hence $I = 0$. In the case of manufacturing facilities $n \in N^J$, only production and outbound flow occur; therefore, the material balance can be written as:

$$0 = P_{naip} - Q^{OUT}_{naip} \quad \forall n \in N^J, a, i, p \tag{5.29}$$

Where the production term P corresponds of a summation of the batches processed in the F&F process section through lines l, at scale a, in plant $n \in N^J$ and in scenario p.

$$P_{naip} = \sum_{l \in L^S} B_{naslip|s=F} \quad \forall n \in N^J, a, i, p \tag{5.30}$$

The outbound flow for manufacturing corresponds to a summation of Q from nodes $n \in N^J$ to storage nodes $n \in N^K$.

$$Q^{OUT}_{naip} = \sum_{m \in M^N} Q_{tnmaip} \quad \forall n \in N^J, a, i, p \tag{5.31}$$

Furthermore within a manufacturing plant a set of additional mass balances is written to regulate the flow of batches between process sections s. Hence:

$$J_{nasip} = \sum_{l \in L^S} B_{naslip} - \sum_{l \in L^S} B_{na,s+1,lip} \quad \forall n \in N^J, a, s, i, p \tag{5.32}$$

Where J corresponds to the inventory of batches held between process section s upstream and its downstream process section $s+1$. This holds true for upstream to downstream process section and downstream to F&F parts of the process. In this study, a single-period problem is considered and no pooling is allowed, hence $J = 0$. Hence:

$$0 = \sum_{l \in L^S} B_{naslip} - \sum_{l \in L^S} B_{na,s+1,lip} \quad \forall n \in N^J, a, s, i, p \tag{5.33}$$

In the case of storage facilities $n \in N^K$, material is not produced or consumed and the expression in Eq. 4.25 simplifies to a balance between inbound and outbound flows:

$$0 = Q^{IN}_{naip} - Q^{OUT}_{naip} \quad \forall n \in N^K, a, i, p \tag{5.34}$$

Where the inbound flows Q^{IN} of storage correspond to a summation of Q from nodes $n \in N^J$ to the storage node $n \in N^K$:

$$Q^{IN}_{maip} = \sum_{n \in M^N} Q_{nmaip} \quad \forall m \in N^K, a, i, p \tag{5.35}$$

Conversely, the outbound flows Q^{OUT} of storage correspond to a summation of Q the node $n \in N^K$ to demand zones $n \in N^Z$.

$$Q^{OUT}_{naip} = \sum_{m \in M^N} Q_{nmaip} \quad \forall n \in N^K, a, i, p \tag{5.36}$$

Finally, in the case of demand zones $n \in N^Z$, the overall material balance in Eq. 5.28 simplified to a balance between inbound flows and product sales as follows:

$$0 = Q^{IN}_{naip} - S_{naip} \quad \forall n \in N^Z, a, i, p \tag{5.37}$$

Where Q^{IN} is computed a summation of Q from upstream storage nodes $m \in N^K$ to each demand zone $n \in N^Z$:

$$Q^{IN}_{maip} = \sum_{n \in N^M} Q_{nmaip} \quad \forall m \in N^Z, a, i, p \tag{5.38}$$

The amount of doses for product i supplied to each demand zone corresponds to the amount of batches S at demand zone $n \in N^Z$ produced with scale a multiplied by the batch size parameter χ. This quantity in doses has to be greater than or equal that the demand D on the R.H.S at the demand zone $n \in N^Z$.

$$\sum_{a \in A} S_{naip} \chi_{aip} \geq D_{ni} \quad \forall n \in N^Z, i, p \tag{5.39}$$

In this study, the demand D is calculated by the following equation based on the total demand for each product D^T_i:

$$D_{ni} = D^T_i / \mathrm{card}(N^Z) \quad \forall n \in N^Z, i \tag{5.40}$$

5.2.3.9 Shortage terms

The sales-demand constraint in the deterministic and stochastic optimisation steps can be modified to account for the possibility of the sales not meeting the target demand for certain uncertainty realisations. This can help avoid infeasible solutions, when the capacity available is not sufficient to cater for the target demand. The shortage variable G is introduced to account for potential shortages at each demand zone $n \in N^Z$, for product i and scenario p.

$$\sum_{a \in A} S_{naip} \chi_{aip} + G_{nip} \geq D_{ni} \quad \forall n \in N^Z, i, p \tag{5.41}$$

This requires a modification in the objective function with the introduction of a penalty term for supply shortage.

$$z = \sum_{p \in P} \phi_p (TC_p^{CAPj} + TC_p^{OPj} + TC_p^{OPk} + TC_p^T + \psi \sum_{i \in I} \sum_{n \in N^Z} G_{inp}) \tag{5.42}$$

Through this reformulation, the objective of the problem is to minimise costs and shortage where possible with respect to both products i. The ψ parameter is selected to be large enough to promote maximisation of supply and avoid solutions where it is less costly to create shortages.

5.2.3.10 Chance-constraint formulation

The previous formulation corresponds to an optimisation problem where a cost objective function is minimised subject to a set of equality constraints and inequality constraints. The key inequality constraint g in the supply chain formulation is the sales-demand constraints for each of the manufactured products.

The introduction of chance constraints Charnes and Cooper (1959) requires the selection of inequality constraints which are allowed to be satisfied to a predefined probability level. Equation 5.43 illustrates the mathematical formulation of a generic chance constraint:

$$\mathbf{P}[g(\mathbf{x}, \xi \leq 0)] \geq \epsilon \tag{5.43}$$

Where the inequality constraints g have to be satisfied with a probability P of at least ϵ in order for a solution to be feasible. Individual chance constraints can be introduced as follows: for each of the inequality constraints g_1, g_2..., g_n:

$$\mathbf{P}[g_n(\mathbf{x}, \xi \leq 0)] \geq \epsilon_n \quad \forall n \tag{5.44}$$

In the formulated, single-period campaign planning problem, the key inequality is the sales-demand constraints (Eq. 5.39), which quantifies the number of doses supplied by multiplying the number of batches S to the batch size parameter χ. The batch size is an uncertain parameter characterised by a probability distribution for each scale a and each product i. The number of sales also ties to the time allocation constraints through equality constraints (Eq.5.25, 5.27) which also comprise uncertain parameters for α and cycle time τ. Similarly, cost coefficients c are also scenario dependent and are connected via equality constraints.

The following chance constraint is introduced in the planning model allow the specification *a priori* of the target reliability of the solution generated by the optimisation problem.

$$\mathbf{P}[(\sum_{a} S^L_{naip}\chi_{aip} \geq D_{ni})] \geq \epsilon_i \quad \forall i \tag{5.45}$$

Specifically, we introduce a scenario-based approximation of the chance constraint, which relies on approximating the probability distribution of uncertain parameters ξ by a discrete set of samples P, therefore writing a constraint for each sample and introducing p set of big-M constraints:

$$D_{ni} - \sum_{a} S^L_{naip}\chi_{aip} \leq (1 - y_{p,i})M \quad \forall n \in N^Z, i, p \tag{5.46}$$

Where $y_{p,i}$ are binary variables that mark the activation of the constraint for each sample p and M is an arbitrary large number. If $y_{pi} = 1$ the constraint is active with respect to product i, otherwise it is violated. Finally by imposing the following inequality we effectively approximate the probability of satisfying the chance constraint for each of the products.

$$\sum_{p} \phi_p y_{pi} \geq \epsilon_i \quad \forall i \tag{5.47}$$

5.2.4 Stochastic simulation framework

The performance of each optimisation solution can be tested via stochastic simulations sampling in $d = A \times I$ following the same methodology illustrated in

Chapter 4 and monitoring the value of δ for each product i is compared to the ϵ_i specified *a priori* to assess potential chance constraints violations.

5.2.5 Case study description

The case study considers manufacturers producing lentiviral vectors for gene therapy applications (LV) and vaccines (AdV) in multi-product facilities. The input parameters for the optimisation are presented in Appendix C. In this study, we consider the following target demands for each product:

1. LV demand (D^T_{LV}): 10,000 doses per year
2. AdV demand (D^T_{AdV}): 400,000,000 doses per year

In this Chapter, the optimisation problem is solved to construct a feasibility region by iteratively solving for discrete realisations of the input probability parameters ϵ_{AdV} and ϵ_{LV}. In the *unconstrained* case, the network structure-related decision variables are free and determined as an output of the optimisation. This approach helps understand the required capacity needed to meet target demands D^T_i for each product i with a prescribed probability level ϵ_i. In the *constrained* case, network structure is fixed *a priori* and the feasibility problem is solved to explore the flexibility and evaluate the maximum reliability of a design with respect to the two competing products i. We evaluate the network structure that corresponds to optimal single-product production for LV applications. The problem is constrained via network design decisions as well as a cost upper bound, thus capturing the limited budget that may be available when retrofitting existing facilities to cater for unforeseen production of a secondary product (AdV).

5.2.6 Computational resources

The framework is built in Python 3.11 implementing the PYOMO package. The MILP is solved with GUROBI v11 on a Intel(R) Core(TM) i7-10700 CPU @ 2.90GHz machine with 64.0 GB RAM. Model statistics are displayed in Table 5.1. The multi-product formulation comprises 87391 constraints, 32986 integer variables (23706 binaries) and 32399 continuous and this holds for all the interations to explore the feasibility space. To address slow convergence resulting from the introduction of the scenario-based big-M constraints the optimality gap

threshold is set to 1%. This reduces CPUs to $< 300s$ for each optimisation solved to construct the Pareto contour plots. It was found that a default of 0% gap lead to termination due to memory usage when solving the entire $\epsilon_{LV} \times \epsilon_{AdV}$. To tackle these scalability issues, one could consider the use of cluster computing platforms, however there is no guarantee that convergence would be faster.

Table 5.1: Solution times required to obtain solution for different combinations of $\epsilon_{LV} \times \epsilon_{AdV}$ for the constrained and unconstrained case

Unconstrained			Constrained		
ϵ_{LV}	ϵ_{AdV}	Time (s)	ϵ_{LV}	ϵ_{AdV}	Time (s)
0	0	3	0	0	4
0.25	0	72	0.25	0	5
0.5	0	157	0.5	0	6
0.75	0	44	0.75	0	8
1	0	98	1	0	4
0	0.25	48	0	0.25	5
0.25	0.25	72	0.25	0.25	12
0.5	0.25	166	0.5	0.25	19
0.75	0.25	275	0.75	0.25	14
1	0.25	141	1	0.25	5 (infeasible)
0	0.5	18	0	0.5	7
0.25	0.5	252	0.25	0.5	16
0.5	0.5	84	0.5	0.5	8
0.75	0.5	181	0.75	0.5	16
1	0.5	181	1	0.5	4 (infeasible)
0	0.75	17	0	0.75	7
0.25	0.75	168	0.25	0.75	14 (infeasible)
0.5	0.75	156	0.5	0.75	22 (infeasible)
0.75	0.75	162	0.75	0.75	9 (infeasible)
1	0.75	63	1	0.75	5 (infeasible)
0	1	18	0	1	4 (infeasible)
0.25	1	77	0.25	1	3 (infeasible)
0.5	1	109	0.5	1	4 (infeasible)
0.75	1	210	0.75	1	4 (infeasible)
1	1	29	1	1	5 (infeasible)

The simulation framework helps gain an understanding of solution accuracy. Hence, satisfactory simulation performance demonstrates that not reaching global optimality still leads to good candidate solutions. A single simulation run requires CPUs of 1s$\times K$, which in this case corresponds to $K = 2^6$. In this study, the number of runs is set to 100 and convergence with respect to the standard error (S.E.$\leq 5\%$) is checked after termination of the algorithm. Since each run is independent the procedure could be parallelised, reducing end-to-end computational times and allowing for larger K values.

5.3 Results & discussion

5.3.1 Optimisation

5.3.1.1 Unconstrained case

The presented optimisation framework can be used to optimise network structure so that the reliability of the network is satisfied with respect to the two products. By specifying levels of probability of demand satisfaction ϵ_{LV} and ϵ_{AdV} within the individual chance constraints, a range of structures can be generated which correspond to values of capital and operating costs within the objective function.

The analyses focus only on the manufacturing, as storage and transport costs are not found to be key drivers in the supply chain cost. Storage is considered to be centralised and transport cost is minimised which results in Plant 1 being often the first facility installed in the network as its location is at a closer distance to the centralised storage, followed by Plant 2 and Plant 3. As can be seen in Figure 5.3, network capacity is installed in this order and this is a result of the minimisation of transport cost where possible.

Figure 5.4 displays a contour plot which records (a) operating cost, (b) capital investment as a function of levels ϵ_i of chance constraint satisfaction for each product. Each square within the plot corresponds to a scenario-independent supply chain design which minimises the expected costs. In Figure 5.3, instances of the structures underlying the contour maps are displayed, alongside the expected production and transportation flows which are averaged and normalised over all process scenarios p.

Firstly, one could focus on the results obtained by effectively solving a single-

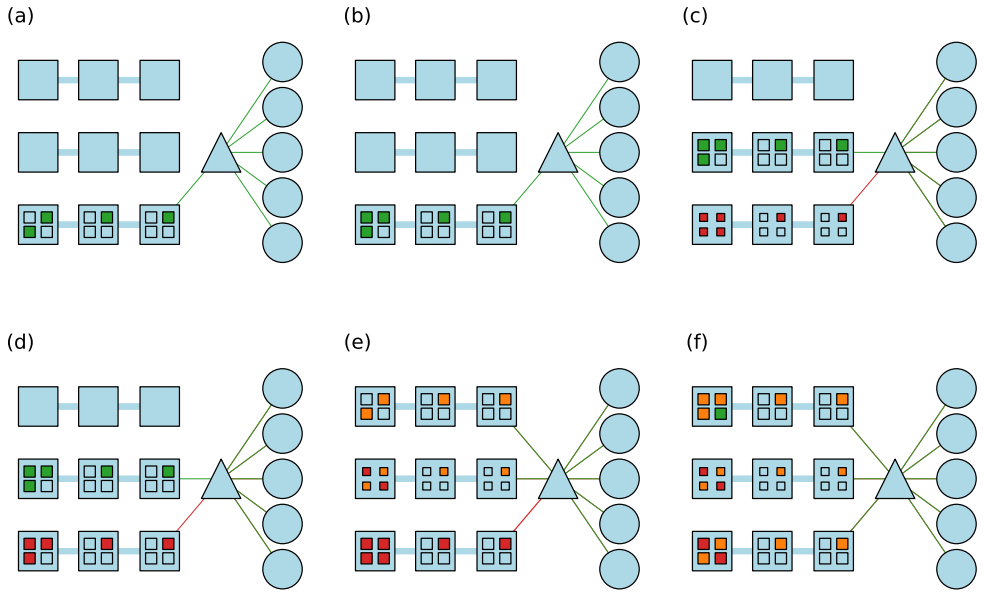


Figure 5.3: Networks for unconstrained case (a-f): AdV production (■), LV production (■), multi-product production, (■), uninstalled inactive lines (□).

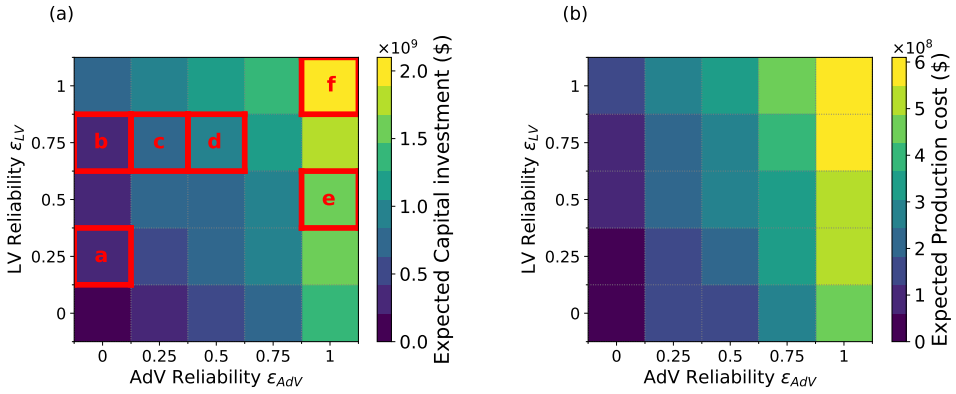


Figure 5.4: Cost performance unconstrained case: (a) capital cost, (b) operating cost

product planning problem for a range of ϵ_{LV} values, whilst $\epsilon_{AdV} = 0$. Imposing the latter condition, allows the optimisation to find a solution that results in no cost and manufacturing resources allocated for the AdV product. Hence, the structures obtained within the square labelled *a*-*b* on Figure 5.4 correspond to networks producing solely LV product for different levels of reliability.

For $\epsilon_{LV} = 0.25$, Network *a* (Figure 5.3a) manufactures relying on 3 upstream reactors operating at 2000L scale, connected to a single manufacturing line in DSP and F&F parts of the process. As $\epsilon_{LV} = 0.75$, Network *b* (Figure 5.3b)

is established, which corresponds to a structure that has a probability of 75% to meet the LV target demand. To do that, it requires the establishment of an additional set of 3 parallel USP lines in Plant 1, resulting in an increase in both capital cost (Figure 5.4a-b).

Network c (Figure 5.3c) corresponds to a structure requiring additional investment in 3 USP parallel lines at 1000L in a third facility, resulting in larger capital cost and operating cost(Figure 5.4a-b) to supply LV with a reliability of $\epsilon_{LV} = 0.75$ and $\epsilon_{AdV} = 0.25$. Cost increase in Network d too, which requires scale which requires scale up of Plant 1 to 3 USP lines at 2000L to supply the target demand with a reliability of $\epsilon_{AdV} = 0.5$. A near-equivalent cost is obtained with Network e, which is obtained by solving an optimisation for $\epsilon_{LV} = 0.5$ and $\epsilon_{AdV} = 1$. Finally, the largest cost results from solving a problem where $\epsilon_{LV} = 1$ and $\epsilon_{AdV} = 1$. In this case, the network operates facilities at both 2000L and 1000L scale, with a number of change-overs which suggest that the network is operating at near full-capacity and is maximising asset utilisation.

Overall, as the target reliabilities ϵ_{LV} and ϵ_{AdV} increase the cost of production and also the capital investment increases. This trend is driven by the fact that the optimisation is constrained to mitigate against an increasing set of scenario realisations the first optimisation solutions will activate scenarios of best case performance (i.e. larger batch size, shorter production times) via the binary variables y_{pi}. Progressively, as the reliability increases the optimisation is constrained to account for the worst-performing scenarios too, thereby requiring larger capacities and costs to supply to an increasing level of reliability.

Interestingly, the results in Figure 5.4a-b show that for different combinations of ϵ_{LV} and ϵ_{AdV} one can find solutions that correspond to the nearly same value of the objective function. Each *iso-cost* solution corresponds to a different network structure in this case. An instance of *iso-cost* solutions is the set of Networks c, d and e, which are also displayed in Figure 5.3c-e.

Through this exercise, decision-makers in the sector could gain an understanding of the required capacity to meet demands for the two competing products at a pre-specified level and simultaneously. The quantification of the required capacity is crucial at the supply chain design stage, where capacity plans can be made based on knowledge of the probabilities of meeting demands of interest and uncertainty awareness. As argued in Chapter 4 there could be a variability in the design decisions for the stochastic formulated which the results

are based on. However, overall costs are expected to be preserved alongside the trade-off with reliability. This limitation on the guarantee on the uniqueness of design however gives scope for the analyses presented in the next section which explore the feasibility space for a design fixed *a priori*.

5.3.1.2 Constrained case

The previous section has highlighted how one can use the framework to obtain optimal designs that ensure a target reliability level with respect to products competing for the same resources. Similarly, the tool can be used to obtain information regarding the maximum capabilities of a fixed design given a cost upper bound. This bound is introduced to capture industrial scenarios where the budget for retrofitting an existing design is limited. In this fashion, the optimisation framework helps identify *iso-cost* solutions for given fixed scales through the networks.

A network structure is specified *a priori* by fixing *here-and-now* decisions. In this case, the underlying network design corresponds to the structure obtained in Network *g* (Figure 5.5g). This corresponds to Plant 1 operating at 2000L capacity with 3 active upstream manufacturing suites and Plant 2 operating 4 upstream lines in parallel at 1000L capacity. The equipment installation variable M with $i = AdV$ is left free, to allow the optimisation to allocate production of AdV if capacity is available and meet the target demands.

Effectively, in this case resource constraints are introduced in the optimisation, which are either monetary constraints (cost upper bound) or technical constraints (fixed maximum capacity). This helps delineate a feasible solution space (Figure 5.6c-d), where (i) the chance constraints in the underlying solutions are satisfied, (ii) the solution implemented does not exceed the cost upper bounds and (iii) the solutions are effectively the nearest feasible ones to the real $iso - cost$ solutions. Beyond this region, certain combinations of ϵ_{LV} and ϵ_{AdV} result in infeasible solutions which are recorded and displayed in grey on the contour maps.

Within the feasible solution space, the costs increase as ϵ_{LV} and ϵ_{AdV} increase up to a boundary which comprises solutions of Networks *g-l* (Figure 5.6-5.5g-l). Network *h* (Figure 5.5h) corresponds to a solution that should meet 75% of the target LV demand and meet 25% of the AdV vaccine demand. Given the upper

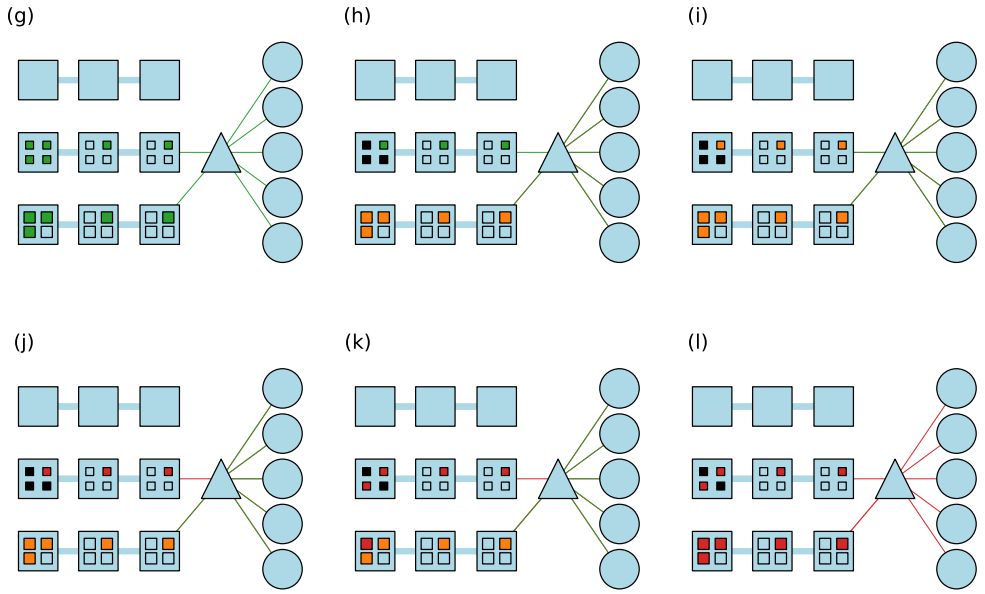


Figure 5.5: Network structures for constraints case (g-l): AdV production (■), LV production (■), multi-product production, (■), installed inactive lines (■), uninstalled inactive lines (□).

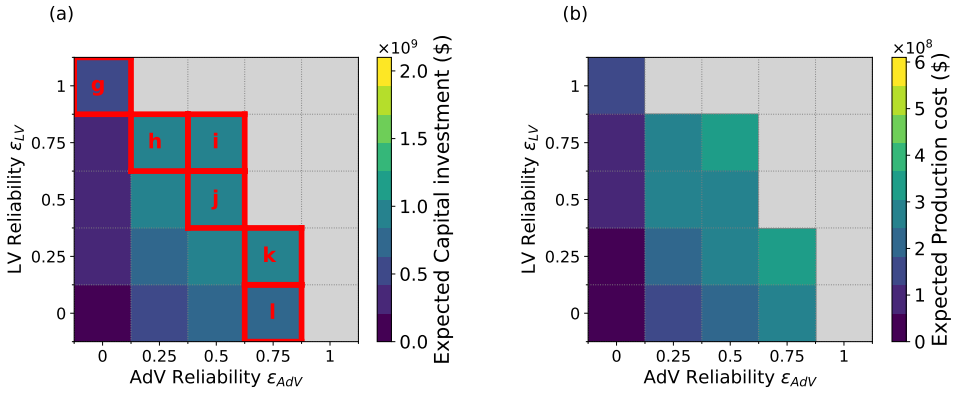


Figure 5.6: Cost performance constrained case: (a) capital cost, (b) operating cost

bound in cost, production of LV is stopped in Plant 2 and the demand on the AdV side is covered by performing a change-over in Plant 1. Increasing ϵ_{AdV} to 0.5, forces the optimisation in Network i (Figure 5.5i) to allocate further AdV production to a second production line in Plant 1 and perform a change-over. Possibly, the facility is operating at near full capacity, in fact a further increase of ϵ_{AdV} to 0.75 results in the next feasible *iso-cost* solution to be manufacturing less LV ($\epsilon_{LV} = 0.25$) in Network j and Network k. Finally, if all resources are repurposed for the manufacturing of AdV ($\epsilon_{AdV} = 0.75$, $\epsilon_{LV} = 0$), Network l

(Figure 5.5l) is obtained.

This highlights that there is a clear competition between products and a trade-off between the achievable production targets at constant capacity and cost. This finding is significant because it informs decision-makers on a number of aspects. Implementing network structures which lie on the boundary (e.g. Network *h-k*) corresponds to operating multi-product facilities which prioritise LV and AdV differently. The back calculation of the corresponding ϵ_{LV} and ϵ_{AdV} informs on the probability of meeting the target demands D^T_{LV} and D^T_{AdV}, which informs on the overall reliability of the system implemented. If this boundary was to move to the right, thereby enlarging the feasible region, this would mean that multi-product systems with larger probabilities to meet the target demands would be achieved. A larger feasible region could be achieved via larger underlying capacities or installing additional facilities or by relaxing the cost-upper bound. The latter would correspond to allowing the network to use additional cost resources to manufacture more and achieve larger production targets and maximum ϵ_i.

Furthermore, the network structure and optimisation output for $\epsilon_{LV} = 0$ would yield the maximum production levels achieved through re-purposing capacity and cost resources for vaccine production entirely. The value of ϵ_{AdV} corresponds to 0.75 in this case study and corresponds to the maximum reliability of the system, which would be achieved if vaccine production was given full priority. If this value was less than 0.75, this would be crucial to inform manufacturers on the attainable production levels and their probability. Based on the obtained probabilities, decision-makers could assess alternative network structures and repeat the analysis through the presented framework.

5.3.2 Simulation

The construction of the cost-resilience plots in Figure 5.4 and 5.6 requires an approximation of the uncertainty space through sampling and the solution of a stochastic optimisation problem. We present the results of a simulation step, which aims at quantifying the accuracy of the optimisation solutions and check whether chance constraints are satisfied at the probability levels ϵ_{LV} and ϵ_{AdV}.

Figure 5.7 displays the results of a single iteration of the Monte Carlo algorithm. The framework uses the cumulative distribution functions with respect to

the two products to quantify for each Monte Carlo run the proportion of samples that satisfies the target demand D_i^T. The service level σ_r for each product corresponds to 1 minus the probability of the network supplying below the target demand value obtained via the CDF curve. This is then used to compute the mean service level δ over a total of 100 Monte Carlo runs.

In Figure 5.7a, Network *a* is simulated and yields service levels $\sigma_r = 0.42$. As network scale increases simulating the solutions obtained via Network *b* yield $\sigma_r = 0.77$ (Figure 5.7b). In most multi-product scenarios (Figure 5.7b-l) σ_r are found to be larger than ϵ for most cases which highlights that the chance constraints are not violated in these cases.

Ultimately, a variability in the value of σ_r for each product is expected in each Monte Carlo run. Hence, Figure 5.8 is displayed to generate a statistical understanding of the level of chance constraint violation and σ_r.

Figure 5.8a-b highlight that statistically the networks can supply to meet the pre-specified demand targets only for Network *a*, given the the mean value of service levels δ is above the target. This does not hold true for Network *b* and would require a re-assessment of whether this accuracy is a result of the number of scenarios to approximate the uncertainty space.

Overall, the chance constraints result well approximated in Networks *c-f*. We argue that δ being greater than ϵ can be explained by the selected objective function during each simulation, which introduces a penalty for any missed deliveries. Given that Network *a* and *b* may not be operating at full capacity, the design is forced by the optimisation-simulation to adjust operational decisions to minimise the additional cost of shortage which outweighs additional production costs. Therefore, the network is driven to supply a larger number of LV therapies which minimises the difference between sales S and demand D.

Conversely, when ϵ_i approaches 1, the minimisation of shortage is no longer relevant as networks are designed to meet 100% of the demand *a priori*. This can be seen in Figure 5.10g and Figure 5.10l which display how the single-product networks perform well in the constrained cost case too. This trend of *oversupplying* to minimise the shortage is also seen in multi-product scenarios. For instance, Network *c* (Figure 5.8c) and Network *e* (Figure 5.8e) yield production levels that meet the reliability target for LV and AdV respectively. In addition to this, spare capacity is used to minimise the therapy shortage with respect to the other product.

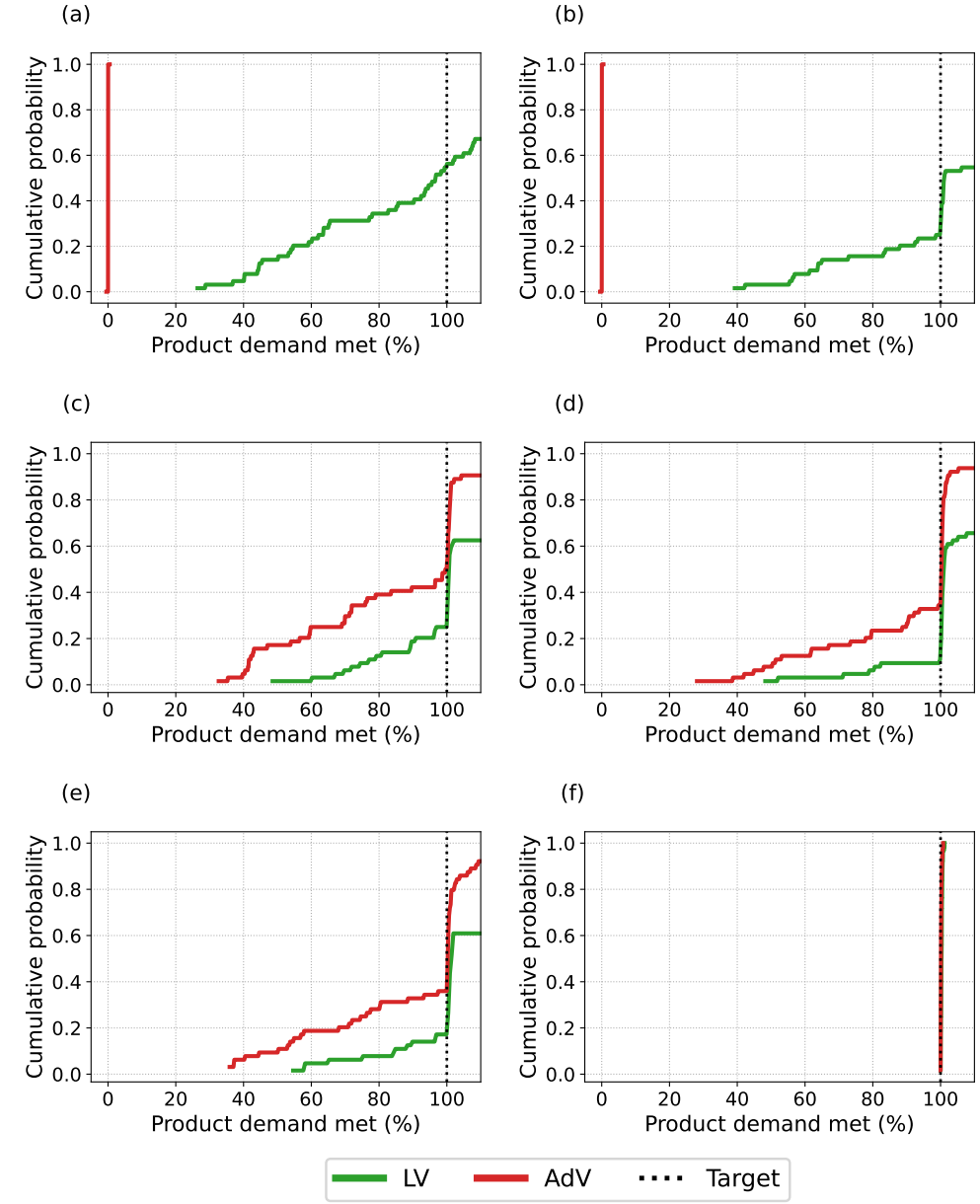


Figure 5.7: Simulations network (a-f), for unconstrained case

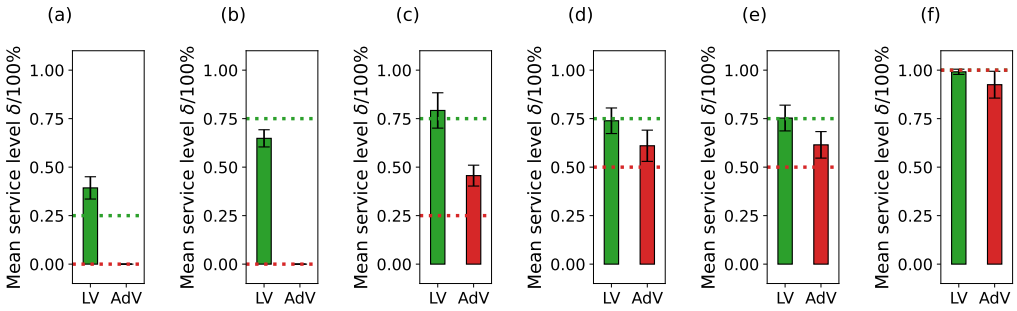


Figure 5.8: Multiple runs simulations for network (a-f), for unconstrained case

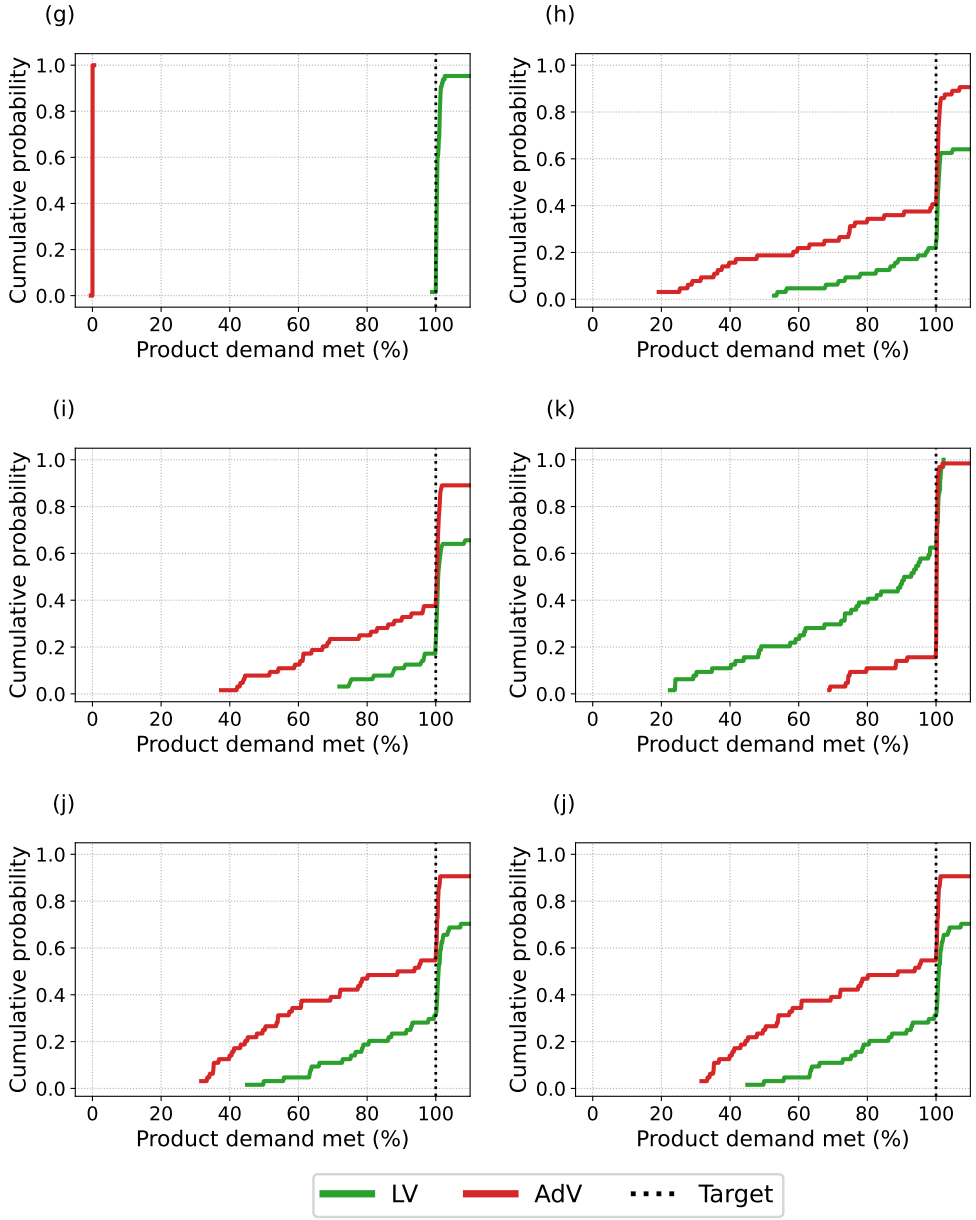


Figure 5.9: Simulations network g to l, for constrained case

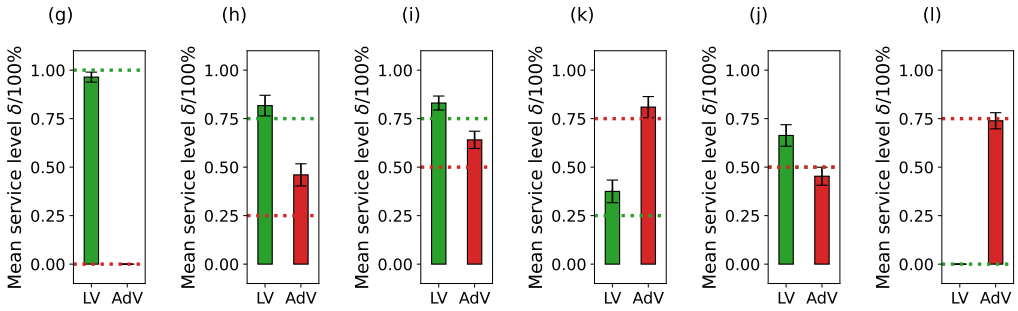


Figure 5.10: Multiple runs simulations for network (g-l), for constrained case

The above observations highlight that chance-constraints are hardly violated, which gives confidence in the generated solutions through the stochastic optimisation. However, in Figure 5.10j Network j performs less accurately in comparison. LV is statistically over the target as opposed to AdV which is statistically under-supplied. This is likely caused by the selection of the shortage cost parameter, which may favour a product over the counterpart. It is expected that a calibration of the penalty cost would be helpful in balancing the mean service level δ with respect to each product (e.g. by scaling by the maximum demand). Furthermore, it is worth highlighting oversupplying and under-supplying being coupled suggests that the chance constraints are well-approximated in the optimisation stage and the simulated results are in fact skewed by the arbitrary value of penalty cost in the objective function upon simulation.

Finally, the trend of oversupplying upon simulation to minimise shortage cost is suggesting that there is an opportunity for solving the problem in smaller intervals of ϵ_{LV} and ϵ_{AdV} and obtain a refined distinction between feasible and infeasible points. This is expected to be a more detailed quantification of the flexibility under uncertainty of a given production asset.

5.4 Conclusions

The decision-support framework presented in this Chapter quantifies the flexibility of production assets under development in catering for multiple products and cope with *unknown* uncertainties. Using stochastic optimisation and chance constraint programming, the framework is used to construct decision-making look up maps for investment planning which support users in identifying the required cost and manufacturing capacity resources to meet desired demands for competing product applications, whilst considering early stage process uncertainty.

We present a case where the probability of demand satisfaction is specified *a priori* and cost-effective network designs are obtained as outputs. It is found that the larger the target reliability, the larger the costs required to ensure satisfaction of the sales-demand constraint. Conversely, we demonstrate the use of the tool to quantify the maximum reliability that a given design and a given expenditure allow. We find a range of *iso-cost* solutions for a given network structure which effectively quantify a feasibility region for the studied supply chain. We demonstrate the accuracy of scenario-based representation of chance constraints

through a Monte Carlo-based stochastic simulation. Furthermore, the need for a more granular resolution of the cost-resilience look up maps is highlighted.

Overall, the constructed cost-resilience maps are a powerful tool for informing investment planning in the current landscape of the next-generation pharmaceutical manufacturing, which leverages on paradigms of flexible production platforms, rapid-scale up and investment planning under process uncertainty. The quantification of combined flexibility and integration of uncertainty are expected to help mitigate impacts of future pandemics. The proposed tool is useful to quantify attainable production targets in case of unforeseen demands which are challenging to anticipate. To this end, it supports the development of flexible platforms which can withstand rapid unforeseen changes in demand.

Once an *unknown* uncertainty is observed in the current time period, the gathered information can be used in investment planning for future time periods, thus becoming a *known uncertainty*. In Chapter 6, we integrate the problem of investment planning with knowledge on process and demand uncertainty.

Chapter 6

Integration of demand uncertainty

Outputs of this chapter have been disseminated in:

- **Sarkis, M.**, Shah, N. & Papathanasiou, M.M.. 'Integrating process and demand uncertainty in capacity planning for next-generation pharmaceutical supply chains', *Submitted to Book of Abstracts of the 35th European Symposium on Computer Aided Process Engineering.*

6.1 Introduction

Investment planning aims at ensuring enough capacity is available to meet a demand forecast. If historical data on previous observed demands is available, one could generate deterministic forecasts which can be met through a series of staged investments. In this sense, deterministic optimisation helps developing a cost-effective investment strategy, under the assumption of perfect information. However, challenges arise when investment planning considers uncertainty in future demands.

For the growing sector of Advanced Therapy Medicinal Products (ATMPs), planning future investments is crucial to ensure its success in making groundbreaking medicines available at all stages from clinical to commercial stages. As argued in Chapter 3, investment planning in this context is challenged by process uncertainty in early stages of decision-making. This may challenge the initial investment decisions, if process capabilities are not known yet. In addition to this, pandemic preparedness is becoming a growing priority and the sector has an opportunity to learn to mitigate against future pandemics by planning capacity expansions early on.

In order to cater for many applications, manufacturers require a careful consideration of different market dynamics and scales. On the one hand, the demand for cancer therapies is quite irregular, with little similarity between subsequent time periods. On the other hand, vaccine demand can rapidly increase in pandemic scenarios or remain equal to zero. If the former outcome realises, demand may prolong for a number of years, rising or declining based on competition and number of other factors. The need to respond to demand fluctuations has seen an increased uptake of single-use equipment in industry to rapidly scale-up and scale-out capacity. Nevertheless, some manufacturers still operate existing stainless steel, which are generally less flexible and require longer set-up timelines for capacity expansion.

The complexity of investment planning problem under demand uncertainty can be represented as a stochastic optimisation and a multi-stage scenario tree for demand outcomes. The PSE literature presents a number of decision support tools based on stochastic programming finding a direct application in capacity planning in pharmaceutical contexts. As discussed in Chapter 2, Rotstein et al. (1999) formulate a two-stage stochastic programming problem considering

clinical trial outcomes uncertainty. First stage variables include initial capacity investments, product selection and initial allocation of manufacturing resources to products. Second stage recourse decisions consist in additional capacity investment or sale of capacity, re-allocation of products to resources. An iterative solution procedure eliminates scenarios with low probability and reduces the computational effort. Gatica et al. (2003a) extend the work by Rotstein et al. (1999) to generate capacity plans under uncertainty over a long time horizon, with a multi-stage stochastic program. With a similar rationale, Tsang et al. (2007a,b) focus on capacity planning under uncertainty for multiple vaccines and extend the work by Gatica et al. (2003a) to evaluate different risk metrics.

The presented literature focuses on formulating multistage stochastic problems to obtain good initial first-stage decisions or initial designs. This serves an initial plan to implement before uncertainty realisation. Once implemented, the multistage planning problem can be solved again at a later time (e.g. at the second stage) and a new set of *here-and-now* decisions is obtained. The abovementioned works do not make a distinction between the use of single- (SU) and multi-use (MU) equipment. The differentiation could yield alternative structural *first-stage* decisions and scenario-dependent *second-stage* decisions based on the recourse actions available, resulting in different costs too. In addition, the discussed frameworks do not consider the integration of process uncertainty in early stages of planning and the subsequent realisation over time. This is because of the assumption that processes have already been intensified and optimised and process capabilities are treated as deterministic.

This Chapter addresses these two emerging knowledge gaps through the development of a multi-stage stochastic planning tool considering SU and MU equipment constraints and the integration of process uncertainty using a deterministic approach. This is done assuming that in a first-stage, decision-makers are risk-averse and would mitigate against worst-case scenarios of process realisations. As time progresses, the true process capabilities become known. The impact of this assumption with respect to process uncertainty is quantified through a comparison with a deterministic case which assumes perfect information of process capabilities throughout the planning.

6.2 Methodology

The framework presented in this Chapter consists of (i) a techno-economic performance quantification step for SU and MU equipment, (ii) optimisation step The following sections illustrate the methodology adopted in the development of the quantification-optimisation framework and discuss the case studies used to illustrate the capabilities of the developed tool.

6.2.1 Problem statement

The problem considered in this study is focuses on capacity planning and is based on the network superstructure illustrated in Figure 6.1. The objective of the planning problem is to minimise total supply chain capital and operating costs and determine the following decisions:

- a network structure and links between nodes;
- investment in manufacturing and storage facilities;
- scales of manufacturing and number of parallel lines;
- scenario-dependent production levels at each manufacturing node;
- scenario-dependent transportation flows between existing nodes;
- scenario-dependent supplied quantities at each demand node.

In the case of SU facilities, the optimisation considers the selection of scale and parallel lines as scenario-depedent (recourse actions). In the case of MU facilities, these become scenario-independent decisions.

To do this, the optimisation considers a range of inputs including:

1. a set T of time periods t;
2. a set I of products i;
3. a set N of nodes n;
4. a set A of manufacturing scales a
5. a set S of manufacturing process sections s;

6. a set L^S of manufacturing process lines l;
7. a set DS of demand scenarios y;
8. datasets of scale-dependent batch sizes, batch times and costs;
9. data on distances between each node n;
10. data on transportation costs;
11. data on storage costs;
12. demand of product i in demand zones $n \in N^Z$ at time t.

The underlying challenge in this investment planning problem is integration of future demand uncertainty and process uncertainty. Whilst demand realises in stages based on the outcomes in each time period, the process uncertainty considered corresponds to datasets quantified through the framework presented in Chapter 3. This corresponds to an uncertainty which is present in early stages of process scale-up and encompasses a range of realisations of future process capabilities. Before these parameters are known, decision-makers may consider adopting a risk-averse approach based on worst-case optimisation in the early stages of the planning horizon. As uncertainty realises over time and/or processes are intensified, true capabilities are known and optimisation decisions can adapt over time. Ultimately, by combining expected demand outcomes and this process uncertainty mitigation strategy, decision-makers can obtain *here-and-now* investment plans which minimise expected costs to implement before uncertainty realisations. Depending on the underlying equipment-related constraints, recourse actions and resulting planning solutions may differ. We discuss the underlying methodology of process performance quantification for different equipment types (SU and MU) in Section 6.2.2 and the formulated multi-stage stochastic problems in Section 6.2.3.

6.2.2 Techno-economic performance quantification

The techno-economic models generated and presented in Chapter 2 are adapted to implement single-use vs stainless steel equipment. This translates preserving the operations and procedure data within each process and amending data regarding the equipment and consumables. Details regarding these modifications

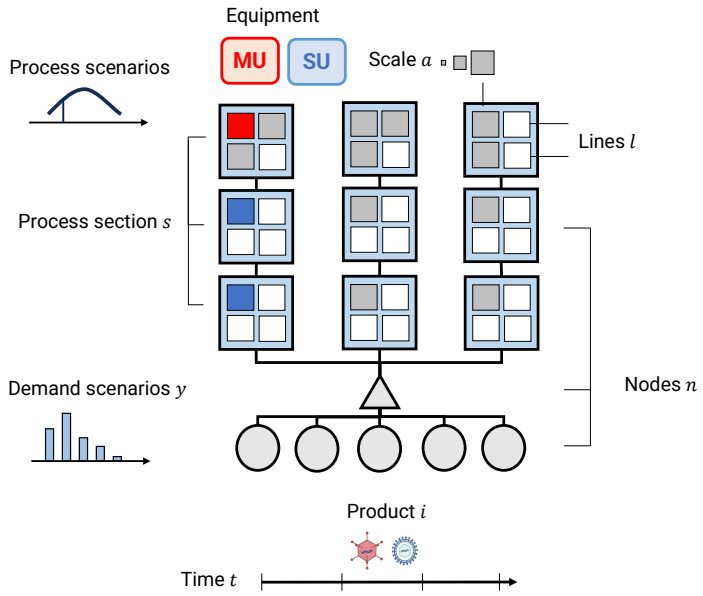


Figure 6.1: Superstructure for optimisation of multi-product i case study. Nodes n comprise manufacturing nodes (square), storage (triangle) and demand zones (circles). Each manufacturing node is modelled as a set of process sections s comprising parallel manufacturing lines l at scale a, according to demand scenarios y at time t.

are provided in Appendix A. In this Chapter, we consider process uncertainty for batch sizes χ, cycle time τ and process time α to be preserved for each scale and product as quantified in Chapter 3. The cost coefficients c are assumed to be deterministic, hence cost uncertainty is not considered for the purposes of the investment planning problem presented in the following section. The full set of inputs for the optimisation is included in Appendix C.

6.2.3 Optimisation formulation

The optimisation considers a multi-period multi-product planning problem where $T = \{T_1, T_2, T_3, T_4\}$ and $I = \{LV, AdV\}$, where LV is a lentivirus-based product for gene therapy applications and AdV is an adenovirus-based product for vaccine applications. In this chapter we adopt a deterministic approach on the process side, considering a single parameter realisation for uncertain parameters at manufacturing level. As explained in Section 6.2.2, throughput-related pa-

rameter are considered uncertain, whereas cost-related parameters are held at the deterministic value quantified via techno-economic modelling. Given that we adopt a staged approach for process uncertainty realisation, process parameters are set to the worst-case scenario value in the first-stage $t = T_1$ and to their median value in stages $t = T_2, T_3, T_4$. On the demand side, a set of demand scenarios y is considered $DS = \{DS_1, DS_2...DS_N\}$, which result from combinations of individual product demands in a multi-stage stochastic scenario tree.

6.2.3.1 Objective function

The objective of the optimisation is to minimise the total expected cost of the supply chain (Eq.6.2). In the SU case, total costs include the demand scenario-dependent capital costs and operating costs of manufacturing, storage costs and transport costs, multiplied by the respective probability Π_y.

$$\min \ z = \sum_{y \in DS} \Pi_y (TC_y^{CAPj} + TC_y^{OPj} + TC_y^{OPk} + TC_y^{T}) \tag{6.1}$$

Where TC^{CAPj} is the total capital manufacturing cost, TC^{OPj} is the total operating manufacturing cost, TC^{OPk} is the total operating storage cost and TC^{T} is the total transport. In the MU case, the capital costs TC^{CAPj} of manufacturing are scenario-independent quantities, as capacity installation is a scenario-independent decision. Hence:

$$\min \ z = TC^{CAPj} + \sum_{y \in DS} \Pi_y (TC_y^{OPj} + TC_y^{OPk} + TC_y^{T}) \tag{6.2}$$

6.2.3.2 Capital manufacturing costs

Capital costs of SU manufacturing (Eq.6.3) are computed as a sum of product-specific equipment investment costs.

$$TC_y^{CAPj} = \sum_{t \in T} \sum_{n \in N^J} C_{tn}^{EQj} \quad \forall y \tag{6.3}$$

The term in the sum of Eq. 6.3 is C^{EQj}, accounting for the total cost of equipment investment at time t, in node $n \in N^J$ and demand scenario y. It is worth noting that in this formulation we not consider product-specific capital costs, because these also lead to product-specific investment constraints. Therefore we

make a simplification and use general line capital cost realisation parameters, which correspond to those for LV.

We introduce a variable of investment decision, H, which marks the investment of a manufacturing line l in process section s and is independent of the product allocated and the demand scenario y.

$$C_{tny}^{EQj} = \sum_{a\in A}\sum_{s\in S}\sum_{l\in L^S} c_{as}^{EQjl} H_{tnalsy} \quad \forall t, n \in N^J, y \tag{6.4}$$

The equipment installation cost C^{EQj} is computed as a function of binaries H, which indicate whether equipment for the manufacturing at scale a is available and installed in lines l of section s in plant $n \in N^J$ at time t in scenario y and capital cost realisations for parameters c^{EQjl}(Eq.6.4):

In the MU case, Equations 6.3 and 6.4 are modified to account for scale-up and scale-out decisions through the binary H to be scenario-independent. Hence:

$$TC^{CAPj} = \sum_{t\in T}\sum_{n\in N^J} C_{tn}^{EQj} \tag{6.5}$$

$$C_{tn}^{EQj} = \sum_{a\in A}\sum_{s\in S}\sum_{l\in L^S} c_{as}^{EQjl} H_{tnals} \quad \forall t, n \in N^J \tag{6.6}$$

6.2.3.3 Operating manufacturing costs

Total SU operating costs TC^{OPj} (Eq. 6.7) of manufacturing correspond to a summation of annual variable costs C^{VARj}, facility related costs C^{FACj} and labour costs C^{LABj} for each manufacturing plant $n \in N^J$, demand scenario y at time t.

$$TC_{y}^{OPj} = \sum_{t\in T}\sum_{n\in N^J} C_{tny}^{LABj} + C_{tny}^{VARj} + C_{tny}^{FACj} \quad \forall y \tag{6.7}$$

Labour costs C^{LABj} for each time period t, manufacturing plant $n \in N^J$ and scenario y are computed for both the SU and MU case (Eq. 6.8) as a summation of the labour costs c^{LABjl} parameters for each scale a, process section l, product i multiplied by batches B respectively which vary with respect to time t, plant $n \in N^J$, scale a, process section s, production line l, product i and demand scenario y:

$$C_{tny}^{LABj} = \sum_{a\in A}\sum_{s\in S}\sum_{l\in L^S}\sum_{i\in I} c_{asi}^{LABjl} B_{tnasliy} \quad \forall t, n \in N^J, y \tag{6.8}$$

Variable costs C^{VARj} for each time period t, manufacturing plant $n \in N^J$ and scenario y are computed for both the SU and MU case (Eq. 6.9) as a multiplication of the variable costs parameters c^{VARjl} for each scale a, product i, process section s by batches B respectively which vary with respect to time t, plant $n \in N^J$, scale a, process section s, production line l, product i and demand scenario y:

$$C_{tny}^{VARj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} \sum_{i \in I} c_{asi}^{VARjl} B_{tnasliy} \quad \forall t, n \in N^J, y \tag{6.9}$$

Facility-dependent SU costs C^{FACj} are tied to depreciation and equipment maintenance costs. In Eq. 6.10, facility-dependent costs for every manufacturing plant $n \in N^J$, time t and demand scenario y are computed as multiplication of facility costs parameters c^{FACjl} for each scale a, process section l and product i by equipment installation-related binaries M respectively which vary with respect to time t, plant j, scale a, process section s, production line l, product i in scenario y.

$$C_{tny}^{FACj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} \sum_{i \in I} c_{asiy}^{FACjl} M_{tnalsiy} \quad \forall t, n \in N^J, y \tag{6.10}$$

In MU cases facility depedent costs are scenario-independent, as the product-specific investment variable M is a first-stage decision. Hence, Equation 6.10 is modified to:

$$C_{tn}^{FACj} = \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} \sum_{i \in I} c_{asi}^{FACjl} M_{tnalsi} \quad \forall t, n \in N^J \tag{6.11}$$

This translates in a modification of Equation 6.7 to account for scenario-independent facility-related costs C_{tn}^{FACj} to the following constraint:

$$TC_{y}^{OPj} = \sum_{t \in T} \sum_{n \in N^J} C_{tny}^{LABj} + C_{tny}^{VARj} + C_{tn}^{FACj} \quad \forall y \tag{6.12}$$

6.2.3.4 Logistics-related costs

In both the MU and SU case, total operating costs of storage TC^{OPk} are a function of (i) the batches handled in the process Q, which corresponds to the number of batches of scale a of product i shipped from upstream storage node $n \in$

N^K to downstream demand zone $m \in N^Z$ in scenario p, (ii) the cost parameter c^{OPk} for product i at scale a.

$$TC_y^{OPk} = \sum_{t \in T} \sum_{n \in N^K} \sum_{m \in N^Z} \sum_{a \in A} \sum_{i \in I} c_{ai}^{OPk} Q_{tnmaiy} \quad \forall y \tag{6.13}$$

Total transport related costs TC^T for each process realisation p are calculated as a function of (i) the distance traveled ρ_{nm} between upstream nodes n and downstream nodes $m \in M^N$, (ii) the number of batches Q at scale a, product i and demand realisation y shipped from n to $m \in M^N$, (iii) a cost parameter for transport c^T.

$$TC_y^T = c^T \sum_{t \in T} \sum_{n \in N^M} \sum_{m \in N^M} \sum_{a \in A} \sum_{i \in I} \rho_{nm} Q_{tnmaiy} \quad \forall y \tag{6.14}$$

6.2.3.5 Logic constraints

Logic constraints govern the construction of feasible supply chain network structures. First, a link between an upstream node n and downstream node $m \in M^N$ is established only if the upstream node n is used at time t.

$$X_{tnm} \leq Y_{tn} \quad \forall t, n \in N, m \in M^N \tag{6.15}$$

Where X is the binary variable that marks the establishment of a link between nodes n to $\in M^N$ and Y is the binary variable that marks the use of a node in location $n \in N$ at time t.

Sourcing constraints ensure that each node is sourced by at least one route via the binary variables X. In the case of storage nodes $m \in N^K$, the node is used if the binary variable $Y = 1$ (Eq. 6.16). In the case of demand zones $m \in N^Z$ the nodes are sourced at all time, with the R.H.S. of Eq.6.17 equal to 1.

$$\sum_{n \in N^J} X_{tnm} \geq Y_{tm} \quad \forall t, m \in N^K \tag{6.16}$$

$$\sum_{n \in N^K} X_{tnm} \geq 1 \quad \forall t, m \in N^Z \tag{6.17}$$

We consider that at most one scale per plant $n \in N^J$ can be installed. Hence, the following constraint enforces that a scale a is selected only if the manufac-

turing node exists. Furthermore it enforces the selection of only 1 scale of manufacturing, since the L.H.S. summation cannot take values larger than 1. For the MU case, the following constraint is included in the formulation:

$$\sum_{a \in A} V_{tna} \leq Y_{tn} \quad \forall t \in T, n \in N^J \tag{6.18}$$

Where V is the binary variable that marks the selection of scale a in location $n \in N^J$ at time t. In SU cases, this variable is a recourse action which is scenario-dependent. Hence the previous constraint is modified as follows:

$$\sum_{a \in A} V_{tnay} \leq Y_{tn} \quad \forall t \in T, n \in N^J, y \tag{6.19}$$

Similarly, a production line l of MU equipment in process section s at scale a is established only if the corresponding scale is established at that location $n \in N^J$ and time t:

$$Z_{tnasl} \leq V_{tna} \quad \forall t, n \in N^J, a, s, l \tag{6.20}$$

In the SU case both the R.H.S and L.H.S. of Equation 6.20 is scenario-dependent hence:

$$Z_{tnasly} \leq V_{tnay} \quad \forall t, n \in N^J, a, s, l, y \tag{6.21}$$

Where Z takes a value of 1 if a line l is established in process section s, at scale a and time t. Product specific equipment installation can occur only if the manufacturing line is available and installed via the binary variable Z. Therefore in the SU case the constraint takes the form of Equation 6.22. Instead, in the MU case the constraint is written as shown in Equation 6.23

$$M_{tnasliy} \leq Z_{tnasly} \quad \forall t, n \in N^J, a, s, l, i, y \tag{6.22}$$

$$M_{tnasli} \leq Z_{tnasl} \quad \forall t, n \in N^J, a, s, l, i \tag{6.23}$$

Where M marks the product-specific equipment installation for product i, at time t, in node $n \in N^J$, at scale a, process section s and line l in demand scenario y (SU case) or in all scenarios (MU case).

In a similar fashion, product allocation in SU equipment is related to W on the L.H.S. of Equation 6.24 can take value of 1 only if the production line is suitable for production of the product, which is modelled via $M = 1$.

$$W_{tnasliy} \leq M_{tnasliy} \quad \forall t, n \in N^J, a, s, l, i, y \tag{6.24}$$

In the MU case, variable M is scenario-independent. Hence:

$$W_{tnasliy} \leq M_{tnasli} \quad \forall t, n \in N^J, a, s, l, i, y \tag{6.25}$$

In both the SU and MU manufacturing case, if two products are manufactured in the same manufacturing node $n \in N^J$, scale a, process section s, line l and demand scenario y and time r, a change over occurs. Binary variables N are introduced to take a value of 1 if a change-over has occurred or 0 if otherwise.:

$$N_{nasly} \leq \sum_{i \in I} W_{nasliy} - 1 \quad \forall n \in N^J, a, s, l, y \tag{6.26}$$

In this multi-product multi-period formulation, change overs occur throughout the time period t and the sequence of production is not considered. Therefore, the following equations are introduced to account for change-overs occurring only once, assuming that one product (LV) is always produced before the other (AdV).

$$C_{nasliy|i=LV} \leq N_{nasly} \quad \forall n \in N^J, a, s, l, y \tag{6.27}$$

$$C_{nasliy|i=AdV} \leq N_{nasly} - 1 \quad \forall n \in N^J, a, s, l, y \tag{6.28}$$

Effectively, by making change-over variables demand scenario y dependent, some flexibility to repurpose equipment is introduced, whether SU or MU-based, This means the formulation captures a manufacturing capability to flexibly adapt to demand realisations, for instance due to market competition.

6.2.3.6 Investment constraints

A set of constraints is added to constrain the scale up decisions in subsequent time periods. In the MU case, once a facility is installed the scale will remain fixed for the subsequent time periods. Conversely, this does not hold for SU based equipment, where equipment of larger scale can be flexibly installed for

one period to the next. To model this the following constraints are included in the MU formulation:

$$Z_{tnasl} = Z_{t-1,nasl} + H_{tnasl} \quad \forall t|t > 1, n \in N^J, a, s, l \tag{6.29}$$

Where the R.H.S. of Equation 6.29 correspond to the existence of a line l in process section s in plant j and scale a at time $t > 1$. This variable takes a value of 1 if it was activated in the previous time period too (L.H.S.) or installed via the activation of H and takes a value of 0 otherwise. Conversely in the SU case, the investment constraint is relaxed to allow scale to be modified across time periods and demand scenarios.

$$Z_{tnasly} \leq Z_{t-1,nasly} + H_{tnasly} \quad \forall t|t > 1, n \in N^J, a, s, l, y \tag{6.30}$$

6.2.3.7 Transport constraints

Transport constraints regulate the maximum and minimum achievable flows between upstream nodes n and downstream nodes $m \in M^N$ in the supply chain network at time t.

$$Q^{MIN} X_{tnm} \leq \sum_{i \in I} Q_{tnmaiy} \leq Q^{MAX} X_{tnm} \qquad \forall t, n \in N, m \in M^N, y \tag{6.31}$$

Where Q^{MIN} and Q^{MAX} correspond to lower and upper bound parameters for the flow variables Q if the link is established via $X = 1$. Instead, these are enforced to be equal to 0 if $X = 0$.

6.2.3.8 Manufacturing constraints

With a similar rationale to transport constraints, time constraints bound the time allocated to production of product i in each process section s, manufacturing line l, process plant $n \in N^J$ at all times t in each scenario y.

$$T^{MIN} W_{tnasliy} \leq T_{tnasliy} \leq T^{MAX} W_{tnasliy} \quad \forall t, n \in N^J, a, l, s, i, y \tag{6.32}$$

Where the bounds are activated if the corresponding product allocation variable $W = 1$, otherwise the variable T is forced to be equal to 0. Additional KPIs

of manufacturing are the total utilisation of manufacturing capacity, which is computed as a summation of the time T allocated to production of each product i (Eq. 6.33). An overall upper bound can also be imposed via Eq. 6.34.

$$U_{tnasly} = \sum_{i \in I} T_{tnasliy} \quad \forall t, n \in N^J, a, s, l, y \tag{6.33}$$

$$U_{tnasly} \leq U^{MAX} \quad \forall t, n \in N^J, a, s, l, y \tag{6.34}$$

The number of batches B of product i processed in each process section s, process line l, equipment e at scale a in location $n \in N^J$ time t and scenario y are calculated as a function of (i) the time allocation variable T, (ii) demand realisations in each scenario y. Specifically, these correspond to the process time taken to collect the first batch in the campaign α in each process section s and the length of the bottlenecks within each process section τ.

$$B_{tnasliy} = W_{tnasliy} + 1/\tau_{st}(T_{tnasliy} - \alpha_{st} W_{tnasliy} - \kappa C_{tnasiy}) \quad \forall t, n \in N^J, a, s, i, y \tag{6.35}$$

The above expression (Eq. 6.35) is written to compute the number of batches B in each campaign as a sum of the first batch ($W = 1$) and the remaining $B-1$ batches which are obtained within the time slot $T - \alpha$.

6.2.3.9 Network constraints

The overall material balance can be written for every node in the supply chain network can be written as follows.

$$I_{tnaiy} = I_{tnaiy|t-1} + Q^{IN}_{tnaiy} - Q^{OUT}_{tnaiy} + P_{tnaiy} - S_{tnaiy} \quad \forall t, n \in N, a, i, y \tag{6.36}$$

Where I is the inventory of batches at scale a for product i held at time t, in node n in scenario y, Q^{IN} and Q^{OUT} denote the inflow of batches from upstream nodes and outflow to downstream nodes, P denotes the amount of material produced and S denotes the amount of material consumed (i.e. sold). This general form of material balance can be re-written for each node type.

In the case of manufacturing facilities $n \in N^J$, only production and outbound flow occur and no transformation of input material is modelled; therefore, the material balance can be written as:

$$I_{tnaiy} = I_{tnaiy|t-1} + P_{tnaiy} - Q^{OUT}_{tnaiy} \quad \forall t, n \in N^J, a, i, y \tag{6.37}$$

Where the production term P corresponds of a summation of the batches processed in the F&F process section through lines l, at scale a, in plant $n \in N^J$ at time t and in scenario y.

$$P_{tnaiy} = \sum_{l \in L^S} B_{tnasli|s=F} \quad \forall t, n \in N^J, a, i, y \tag{6.38}$$

The outbound flow for manufacturing corresponds to a summation of Q from nodes $n \in N^J$ to storage nodes $n \in N^K$.

$$Q^{OUT}_{tnaiy} = \sum_{m \in N^K} Q_{tnmaiy} \quad \forall t, n \in N^J, a, i, y \tag{6.39}$$

Furthermore within a manufacturing plant a set of additional mass balances is written to regulate the flow of batches between process sections s. Hence:

$$J_{tnasiy} = \sum_{l \in L^S} B_{tnasliy} - \sum_{l \in L^S} B_{tna,s+1,liy} \quad \forall t, n \in N^J, a, s, i, y \tag{6.40}$$

Where J corresponds to the inventory of batches held between process section s upstream and its downstream process section $s+1$. This holds for upstream to downstream process sections and downstream to F&F parts of the process.

In the case of storage facilities $n \in N^K$, material is not produced or consumed and the expression in Eq. 6.36 simplifies to a balance between inbound and outbound flows.

$$0 = Q^{IN}_{tnaiy} - Q^{OUT}_{tnaiy} \quad \forall t, n \in N^K, a, i, y \tag{6.41}$$

Where the inbound flows Q^{IN} of storage correspond to a summation of Q from nodes $n \in N^J$ to the storage node $n \in N^K$:

$$Q^{IN}_{tmaiy} = \sum_{n \in N^J} Q_{tnmaiy} \quad \forall t, m \in N^K, a, i, y \tag{6.42}$$

Conversely, the outbound flows Q^{OUT} of storage correspond to a summation of Q the node $n \in N^K$ to demand zones $n \in N^Z$.

$$Q^{OUT}_{tnaiy} = \sum_{m \in N^Z} Q_{tnmaiy} \quad \forall t, n \in N^K, a, i, y \tag{6.43}$$

Finally, in the case of demand zones $n \in N^Z$, the overall material balance in Eq. 6.36 simplified to a balance between inbound flows and product sales as follows:

$$0 = Q^{IN}_{tnaiy} - S_{tnaiy} \quad \forall t, n \in N^Z, a, i, y \tag{6.44}$$

Where Q^{IN} is computed a summation of Q from upstream storage nodes $m \in N^K$ to each demand zone $n \in N^Z$:

$$Q^{IN}_{tmaiy} = \sum_{n \in N^K} Q_{tnmaiy} \quad \forall t, m \in N^Z, a, i, y \tag{6.45}$$

The amount of doses for product i supplied to each demand zone corresponds to the amount of batches S at demand zone $n \in N^Z$ at time period t produced with scale a multiplied by the batch size parameter χ. This quantity in doses has to be greater than or equal that the demand D on the R.H.S at the demand zone $n \in N^Z$.

$$\sum_{a \in A} S_{tnaiy} \chi_{ai} \geq D_{tniy} \quad \forall t, n \in N^Z, i, y \tag{6.46}$$

In this study, the demand D is calculated by the following equation based on the total demand for each product D^T_i:

$$D_{tniy} = D^T_{tiy} / \text{card}(N^Z) \quad \forall t, n \in N^Z, i, y \tag{6.47}$$

6.2.3.10 Non-anticipativity constraints

Non-anticipativity constraints in multi-stage stochastic programming ensure that decisions made at any given stage rely only on information available at that time, without foreseeing future outcomes. Hence, for every scenario-dependent decision indexed by scenario y, a non-anticipativity constraint is introduced as illustrated in Equation 6.48:

$$x_{t,y_1} = x_{t,y_2} \quad \forall y_1, y_2 \in \Delta^t \tag{6.48}$$

Where x_{t,y_1} and x_{t,y_2} represent all scenario-dependent decisions made at stage/time period t and Δ^t is the set of scenario pairs, which are indistinguishable at stage t. At stage $t = T_1$, Δ^t comprises all scenarios, whilst these branch out in subsets as time periods progress. These constraints ensure that decisions at stage t is consistent across all scenario within Δ^t, reflecting that the decision-maker does not yet know which demand scenario will occur beyond the information available at stage t.

6.2.4 Case studies

6.2.4.1 Planning strategies

Given the datasets generated through techno-economic modelling for a number of scales and gathered in Appendix A, we perform a comparison between a multi-stage stochastic planning problem implementing SU equipment and one implementing MU equipment. As discussed in Section 6.2.3, the implementation of SU or MU leads to different recourse actions upon uncertainty realisation which results in different multi-stage formulations.

We consider a case where process knowledge/process improvements will be registered after a period of time after manufacturing starts. In this manner, one can model a risk-averse decision-making strategy whereby to avoid undersupplying in the first year, one can perform a worst-case optimisation and then assume that processes will improve as in the coming years/knowledge will be acquired. This corresponds to adopting a staged optimisation approach on the process uncertainty side, where confidence on process performance improves with time. In addition, the formulation is kept stochastic only on the demand side. This strategy is compared with a case where perfect process knowledge is assumed from the first time period to monitor how the solution changes in terms of production flows, inventories and expected costs.

6.2.4.2 Demand scenarios

In this Chapter, we consider a scenario tree which is made up of two combined product demand trees, namely one for LV-based cancer therapies and AdV-based vaccines. Overall, the two scenario trees combined result in a total of 216 scenarios $y \in D^S$ each with a computed probability of Π_y.

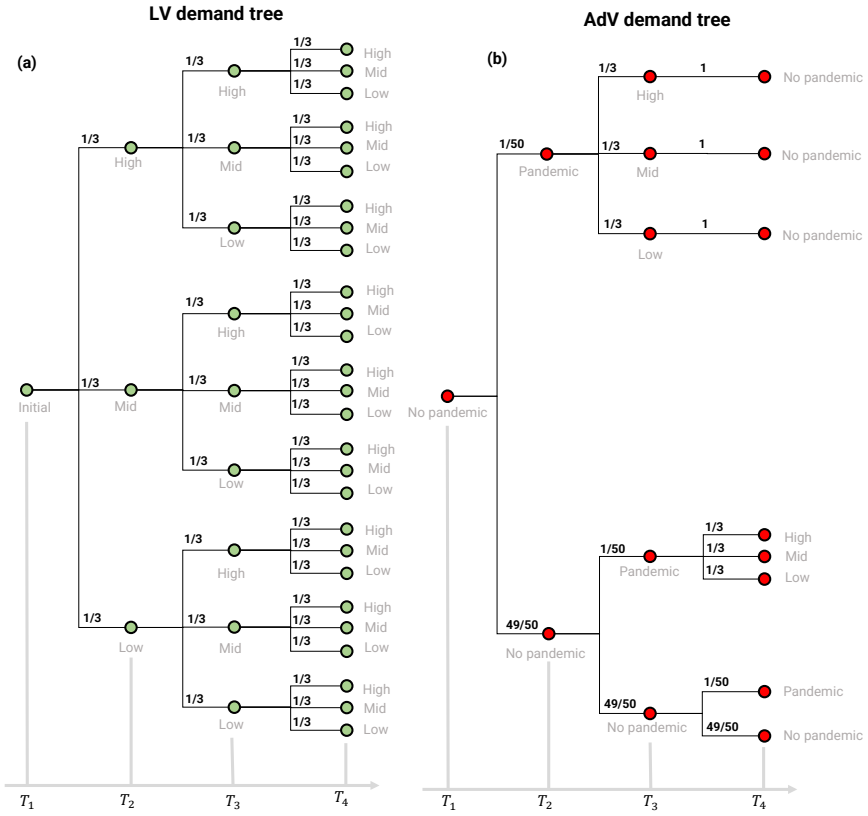


Figure 6.2: Demand scenarios: (a) LV demand tree, (b) AdV demand tree. The optimisation considers a joint scenario tree based on these outcomes for each product.

As can be seen in Figure 6.2a, the LV demand trees consider that at a first stage, before demand uncertainty is realised, demand corresponds to a nominal value (10,000 doses per year). As time progresses, demand may realise to be *high* (20,000 doses per year), *mid* (10,000 doses per year), *low* (5,000 doses per year), with equal probabilities ($\frac{1}{3}$). This also holds for later stages of uncertainty realisation. Through the LV scenario tree, one reconstructs time-dependent combinations of demand uncertainty for cancer therapies, which are characterised by irregular demands and generally hard to predict. The scenario tree could be modified to embed more refined trend based on the information available to decision-maker. Overall, by considering 4 stages of uncertainty realisation one obtains 27 scenarios on the LV side.

Figure 6.2b illustrates how vaccine-related demands may vary through time and aims to model uncertainty in demand through estimated yearly pandemic probabilities. According to a study by Marani et al. (2021), there is a probability of a pandemic outbreak every year of approximately $\frac{1}{50}$. Based on this

Table 6.1: Computational statistics of the snapshot optimisation framework

Model	CPU (s)	Constraints	Integer (Binary)	Continuous	Gap
MU (base)	54,747s	878,916	1,369,544(1,120,280)	878,916	7.82%
MU (worst)	25,715s	878,916	1,369,544(1,120,280)	878,916	7.80%
SU (base)	36,012s	2,258,780	1,379,864(1,130,600)	878916	4.38%
SU (worst)	36,513s	2,258,780	1,379,864(1,130,600)	878,916	4.19%

assumption one can construct scenario trees based on a *non-pandemic* outcome and a *pandemic* outcome each year. In the former case, the scenario tree would branch into a pandemic/non-pandemic in the next stage. In the latter case, the demand for vaccines may be *high* (400 million doses per year), *mid* (200 million doses per year), or *low* (100 million doses per year). The *high* outcome may capture scenarios of novel variants emerging or competition favouring the produced vaccine. The *low* outcome captures scenarios where demand decreases because of more successful competitors and/or issues with side effects post-commercialisation (EMA, 2021). Alternatively, the demand can remain stable over the time horizon *(mid)*. Next, one could assume that the pandemic would resolve in the space of n years. In this Chapter we assume n=2, with demand for AdV becoming 0. This type of modelling results in a total 8 demand scenarios on the AdV side.

6.2.5 Model statistics

The MU multi-stage formulation involves a smaller problem size compared to SU because it involves less scenario-dependent decisions (Table 6.1). The model was implemented in Pyomo and solved on an cluster computing using 1 node, 48nCPUs and a memory limit of 100GB. The optimality gap threshold is set to 10% to obtain a solution to compare within 48h. The slow convergence to the solution could be a result of the size of the problem. We discuss strategies to improve the scalability of the problem in Section 6.3.3.

6.3 Results & discussion

6.3.1 Techno-economic analysis

Techno-economic analysis consists in the first step towards understanding cost trade-offs that emerge supply chain cost optimisation. The outputs of this step constitute the input values for the cost coefficients c in the optimisation, which are defined for each cost category VAR, FAC and LAB, process section s, equipment type e, scale a and product i. Figure 6.3 illustrates the cost breakdown for LV and AdV production platforms at different scales implementing either SU or MU equipment.

Overall it emerges that implementing disposable equipment reduces manufacturing costs, especially facility-related costs. As can be seen in Figure 6.3a-d, cost reductions up to 50% are achieved in the SU case for the production on LV in all parts of the process. The reductions in facility-related costs are less significant in the case of AdV, as facility costs are dominated by F&F equipment costs (Figure 6.3b-h). In the developed model, disposable equipment is only introduced in the formulation step, whereas the remaining machinery remains the same between the SU and the MU process case.

In comparison, labour-related costs are reduced by approximately 10% with respect to labour cost because of resulting shorter process times in the SU process. As mentioned in Chapter 3, an underlying assumption of the model is the linear relation between process times and total labour cost per batch which is obtained by multiplying process time by a predefined hourly rate. Due to shorter process times with removed steam-in-place /clean-in-place (SIP/CIP) steps, SU equipment results in lower labour costs per batch.

Focusing on variable costs which embed cost of materials and consumables, the cost per batch in SU equipment is larger than the MU case (Figure 6.3a-d). This is due to a larger number of consumables being required in SU manufacturing compared to MU. This is also seen in AdV manufacturing (Figure 6.3b-h), where variable costs for USP, DSP and F&F result higher than the MU counterpart. Overall the percentage increase varies due to process differences. These include differences in the materials manufacturing and different numbers of process steps in LV and AdV manufacturing respectively.

Figure 6.3 displays costs normalised per batch, showing that SU results in

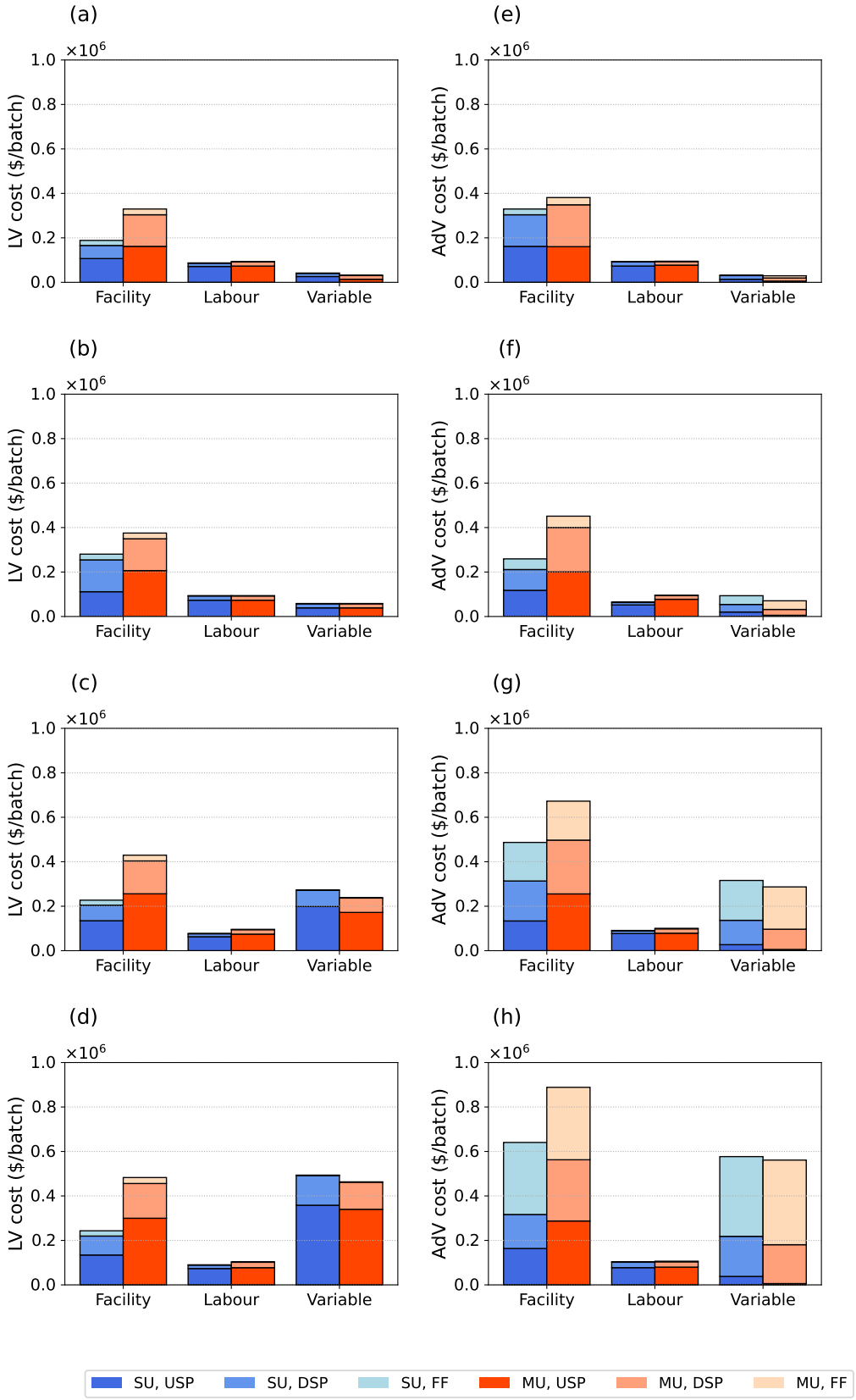


Figure 6.3: Cost breakdown for SU vs MU manufacturing platforms for LV manufacturing at (a) 50L, (b) 200L, (c) 1000L, (d) 2000L scale and AdV manufacturing at (e) 50L, (f) 200L, (g) 1000L and (h) 2000L scale.

lower costs per batch compared to MU. The reductions in facility-dependent cost outweigh the marginal larger variable costs and this is expected to dictate the selection of SU equipment for all process sections, if a cost optimisation is performed.

6.3.2 Investment planning under demand uncertainty

In this analysis we compare an investment planning problem for a manufacturing network implementing MU-based equipment and one based on SU equipment, constraining the optimisation to select either SU or MU equipment. As discussed in the methodology, the two cases consider the same demand scenario tree, however present different recourse actions. In the case of an MU facility, the time delay into setting up new manufacturing lines and facilities requires a plan to be determined a priori with no time for expanding to new lines. When the demand is seen, it takes 2-5 years until a new facility comes in action. In this way, we consider that the *here-and-now* decisions become the 4-year plan for investments, which is scenario-independent. Conversely, in an SU facility additional lines and manufacturing capacity can be installed upon realisation of the demand.

Figure 6.4a-d displays the network scale-up for the MU case, assuming a known process capability in the first year planning. The networks correspond to the solution for scenario $y = Y_1$. As can be seen in Figure 6.4a, the network allocates LV production to two 2000L USP lines and as time progresses it requires a staged investment which is scenario-independent. This translates in an investment which should cover for all scenarios $y \in DS$, including pandemic-related outcomes. Hence, the resulting network structures are over-designed to cater for high demand scenarios of vaccines, although of low probability.

Conversely, the network structures in Figure 6.4e-f correspond to the time-dependent solutions obtained for the SU case for the same demand scenario $y = Y_1$. Given the possibility to rapidly install and de-install manufacturing and amend scale as demand realises the scenario-dependent scale-up results more flexible than the MU counterpart. Specifically, this flexibility to de-install capacity results in lower equipment related cost. The optimisation in the SU manufacturing case is given the degree of freedom of M taking values of 1 or 0 irrespective of the value in the previous time period. Hence, to minimise equipment cost, the networks in Figure 6.4 rely on inventory to meet the demands in the current

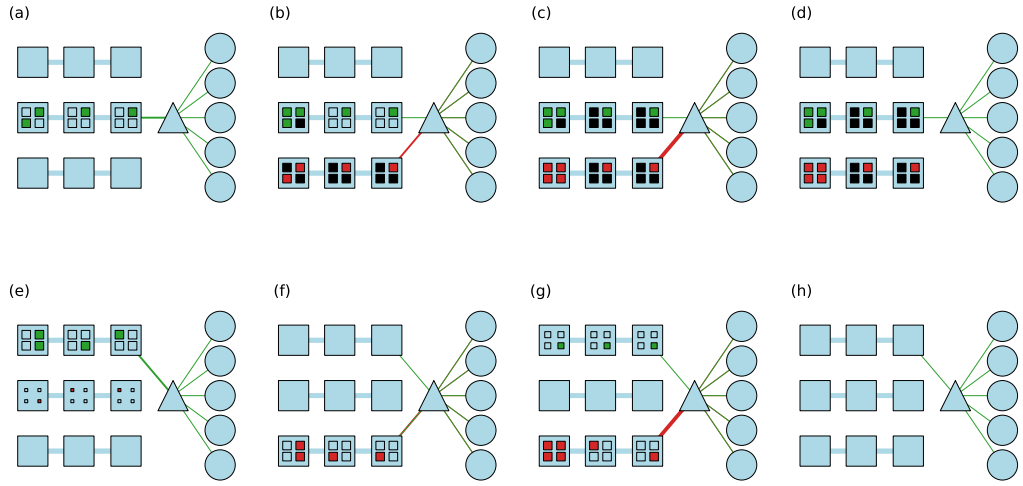


Figure 6.4: Multistage optimisation structures for deterministic case for MU case (a-d) and SU case (e-h). AdV product allocation (■), LV product allocation (■), inactive installed lines (■), inactive uninstalled lines (□)

time periods $t = T_2$ (Figure 6.4f) and $t = T_4$. In this fashion, the reduced manufacturing cost outweighs the cost of storage. This is not an option in the MU case, due to the lifetime of equipment extending beyond the time horizon (20y) and de-installation not being a degree of freedom in the model. Hence, networks in Figure 6.4a-d) involve manufacturing within each time period to supply demands using the cumulatively established capacity, alongside some reliance on accumulated inventories.

In Figure 6.5, overall trends of KPIs in the MU case and SU case are displayed, with a focus on showcasing both the scenario-dependent solutions and the expected values for cost and supply related-KPIs. As can be seen in Figure 6.5a, the USP volumetric capacity in the MU case ends up being scenario-independent, with a starting value of approximately 4,000L bioreactors and scaling to 16,000L in the subsequent time-periods. This scale-up determined a priori as *here-and-now* decisions differs from the volumetric scales obtained in Figure 6.5f. As argued previously, the stochastic optimisation in this case considers the scale-selection and line installment decisions as recourse actions, hence, allowing for a range of scale-up plans to be implemented.

When comparing expected costs of manufacturing displayed in Figures 6.5b-c (MU) and Figure 6.5g-h (SU), it emerges that the expected manufacturing costs in the SU case are lower than the MU case. On the one hand, this trend can be

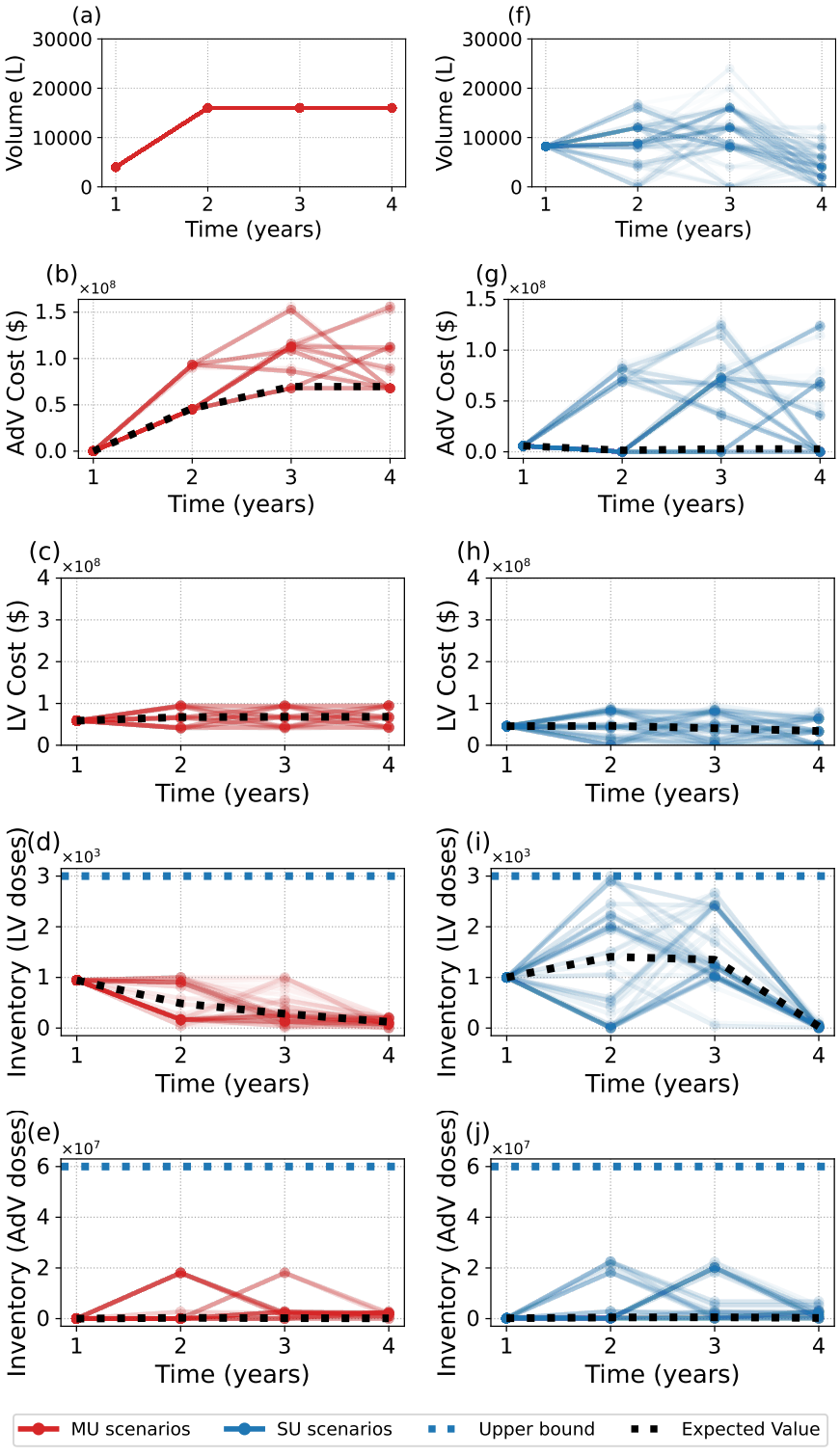


Figure 6.5: Multistage optimisation KPIs for deterministic case for MU case (a-e) and SU case (f-j)

explained by the overall lower costs of manufacturing per batch and the facility-dependent contribution (Figure 6.3). To supply the same demands, one would expect the SU installation to result in lower supply chain costs overall and this is seen in Figure 6.5c and h, where LV expected costs in the SU case remain 70% lower than the MU case. Overall, given LV demand ranges between 5000-20000 doses in all time periods, cumulatively both networks manufacture the same amounts and SU reduces costs further only due to some inventory management. On the other hand, the larger expected costs in MU manufacturing are a result of its inflexibility. Figure 6.5b displays an expected cost of vaccine manufacturing over time for MU equipment, which is larger than the expected cost of SU manufacturing. This is due to the fact that at each stage the optimisation is forced to install capacity to cover for the realisation of a pandemic scenario and resulting implications in later years. Hence, in these cases, repurposing MU equipment via the activation of change-over binary variables C does not provide enough capacity to cater for such high AdV demands. Thus, the facility-dependent cost of AdV manufacturing increases as time progresses so as to cover the *high* scenario following a pandemic outbreak. In time period $t = T_4$, the previously high cost scenario decreases to the lowest possible cost, whereas the pandemic scenario in the third stage may increase again. The expected cost of AdV seems to plateau at a larger value compare to SU case and this is due to the structure of the assumed scenario tree, which models a *high*-demand post-pandemic outbreak scenario in each stage of the scenario tree and forces to *here-and-now* MU decisions to cover for this worst-case outcome.

6.3.3 Integration of process uncertainty

Investment planning in the context of ATMPs is challenged by an inherent process uncertainty in the early stages of decision-making. This uncertainty realises throughout the planning due to a progressive understanding of the attainable production targets or intensification efforts. In this part of the analysis, we focus on a potential strategy for the integration of the uncertainty quantified in Chapter 3. We consider the integration of process uncertainty through a initial worst-case approach, which translates in solving the first stage of decision-making prior to uncertainty realisation with a worst-case approach and considering a deterministic realisation of process uncertainty after a year of operation. Given

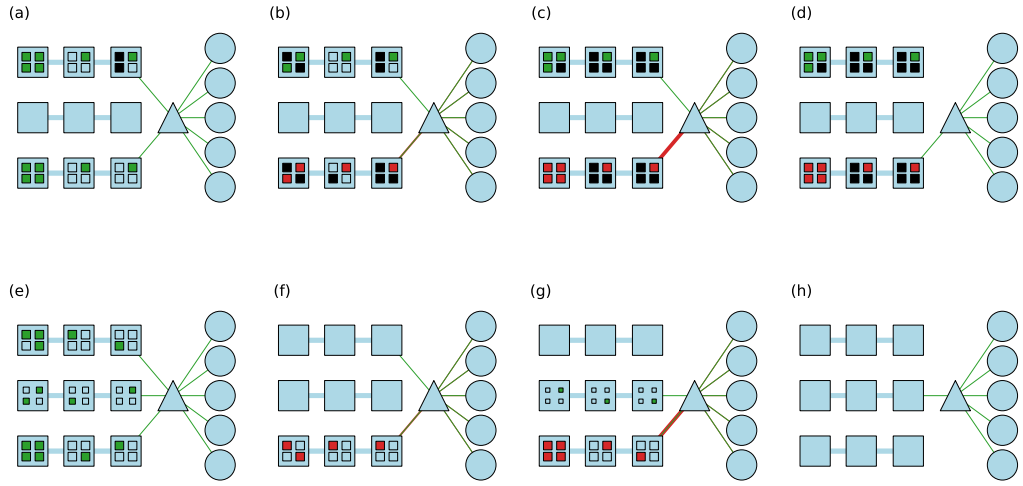


Figure 6.6: Multistage optimisation structures for initial worst-case for MU case (a-d) and SU case (e-h). AdV product allocation (■), LV product allocation (■), inactive installed lines (■), inactive uninstalled lines (□)

that investment decisions are tied to decisions in previous time periods, we investigate whether this risk-averse approach in the early stages of planning has an impact on investment decisions and differentiate results for an MU and an SU case.

In Figure 6.6, we report the network structures obtained through the solution of the investment planning problem with an worst-case approach on the process side. The networks reported correspond again to a single scenario $y = Y_1$ of a total of 216 scenarios in the set D_S. To complement this we also report the KPI trends in Figure 6.7. Overall, the impact of this risk-averse approach to decision-making is mainly seen in first-stage decisions and initial investments. In Figure 6.6a, the optimisation installs a 4 times larger volumetric capacity compared to the initial investment found in 6.4a. This is also seen in the cumulative volume trends in Figure 6.7a. In later time periods, new investments are not required and capacity remains constant, fulfilling all demand scenarios and converging to the same investment plan proposed by the deterministic model (Figure 6.4b-d). Similarly to the deterministic case, the inflexibility of MU equipment leads to an over-design in some scenarios with respect to the target demand, including the displayed scenario Y_1.

Similarly for the SU case, in Figure 6.6e a significantly initial larger capacity is installed compared to the deterministic case (Figure 6.6e). The optimisation

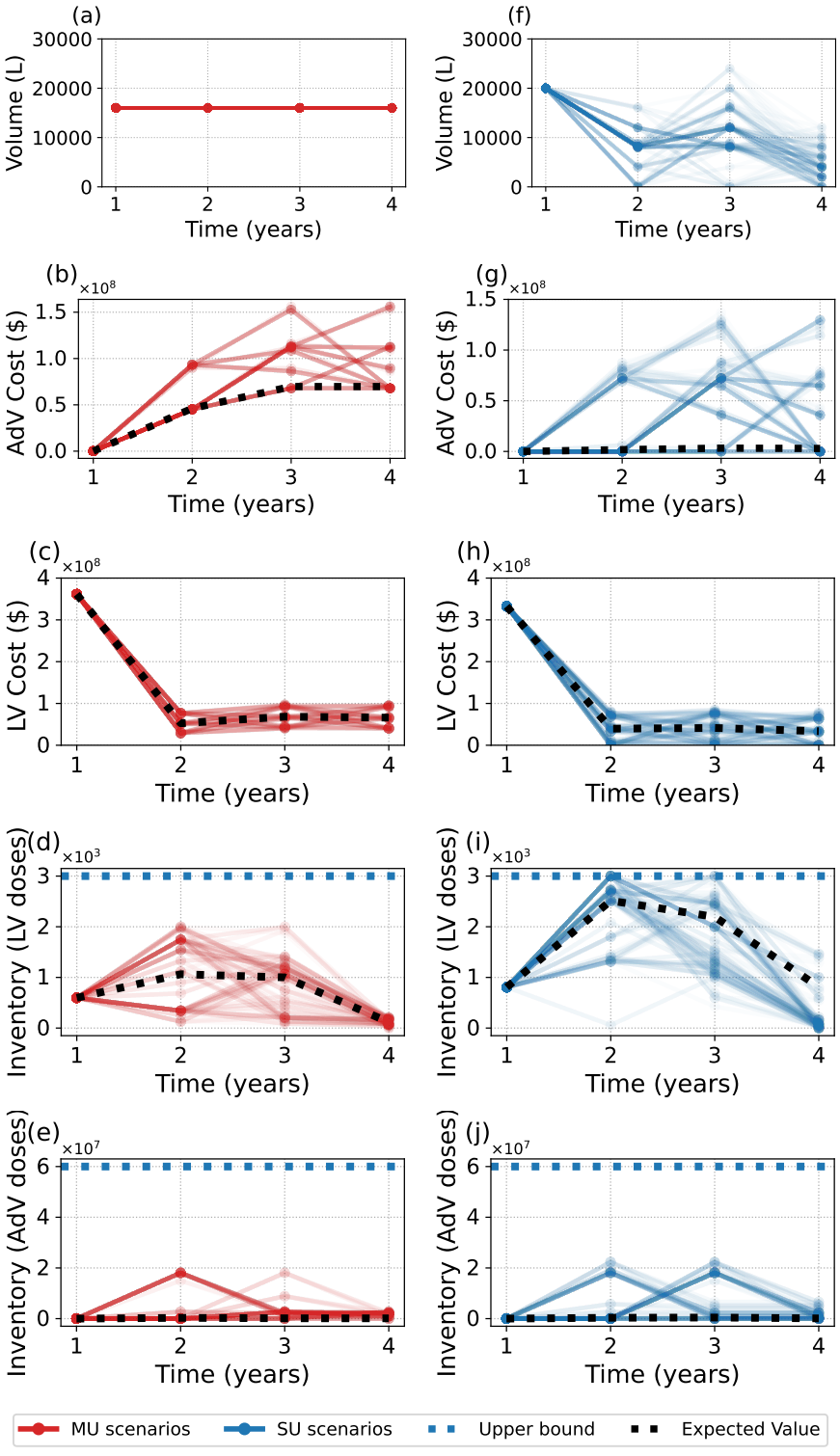


Figure 6.7: Multistage optimisation initial worst-case results for MU case (a-e) and SU case (f-j)

results in structures in time period $t = T_2$ which still receive part of the supply from the inventory and deinstalls LV capacity to save on costs in the second stage. However if we consider the expectation over all scenarios, the larger initial costs of manufacturing contribution in the first stage and lower achievable production levels result in an expected lower inventory accumulated in the first stage (Figure 6.7d) and larger inventories accumulated in later time periods. Effectively the optimisation still reaches an optimal inventory level of approximately 1000 doses to progressively deplete and requiring less manufacturing capacity to be installed.

Interestingly in the case of AdV, the impact of process uncertainty is less significant as through this worst-case approach, the optimisation is solved only in the first stage and with the current scenario tree, AdV demand equals 0 initially. To further develop the model it is worth considering a product-dependent trajectory of the worst-case approach, where process uncertainty is considered to emerge in the first-stage with respect to the first realisation of demand. This would effectively capture how the demand for a new product requires a novel process to be set-up and manufacturing to start before the process is fully intensified and capabilities are known.

In addition, the scalability of the problem is an issue which will require the time horizon of planning to be reduced or decomposition-based approaches to reach solutions faster. If the former route is explored, one could consider solving the full time horizon with a rolling horizon approach, thereby solving at each time step a reduced stochastic problem which considers demand uncertainty and integrates uncertainty in process capabilities with a risk-averse approach once a demand for a new product is realised.

Overall, the current models help develop investment plans which consider demand uncertainty for a range of product applications and integrate considerations regarding the uncertainty in the attainable production targets in early stages of planning. For the first time, we quantify the flexibility to demand uncertainty of SU-based equipment compared to MU equipment and demonstrate that SU promises to lead to lower costs, including cases where process uncertainty is considered in early stages of planning. It is expected that this benefit of adaptability coming form the shorter time scales of recourse actions translates into a better flexibility to unknown unknowns too. Hence, the moving towards SU and modular manufacturing paves the way to more resilient manufacturing networks which can better withstand unforeseen events.

6.4 Conclusion

The pharmaceutical investment planning problem is concerned with making sure capacity is available to meet forecasted product demands. In some instances, demand can be anticipated based on historical data and scenario-trees can be constructed. Typically PSE frameworks for pharmaceutical capacity planning involve a stochatic multi-stage formulation which help *here-and-now* first stage decisions prior to the realisation of uncertainty and represent the initial investment to establish which minimised the expect value of the objective function over a range of scenarios.

In this Chapter, we consider that in the rapidly growing sector of ATMPs initial investments are decided before process capabilities are known. Hence, based on the quantified process uncertainties in Chapter 3 we propose the integration of manufacturing uncertainty in the multi-stage problem through an worst-case approach, whereby process capabilities become known from the second stage of demand realisation and risk-averse decision-making is captured in the initial stage. We compare two multi-stage formulations which capture the recourse actions available in SU facilities and those available in MU equipment and find that the former results in lower expected costs both in a deterministic case and when using worst-case for process uncertainty. This is explained by the shorter time scales for recourse actions of the SU facility, which result in an added degree of freedom to install and remove capacity based on the realised demands. Overall, SU expected costs of planning are lower compared to the MU counterpart, partly due to the added flexibility to install capacity for AdV vaccine manufacturing only in low probability scenarios and an overall cheaper cost of manufacturing due to the use of consumables instead of stainless steel.

For the first time, we demonstrate the added value of adopting SU equipment compared to MU equipment. The models can be further developed to account for different time discretisation of decision-making and introduce scenario-dependent capacity installation with lead times for MU manufacturing. However, the intrinsic shorter timescales of SU manufacturing are expected to consistently result in an advantage to respond to both known and unknown uncertainties. The pharmaceutical industry is already steering towards the use of SU equipment based on the advantage of flexibility. However, a growing concern over the environmental footprint of increased solid plastic waste after each batch is pressuring manu-

facturers to quantify the actual sustainability of SU-based equipment. In the next Chapter, we explore the integration of sustainability metrics in the frameworks proposed in this thesis, illustrate augmented model capabilities and discuss strategies to integrate additional sets of sustainability related uncertainties which ramp up even more the complexity of the investment planning problem.

Chapter 7

Integration of sustainability metrics

Outputs of this chapter have been disseminated in:

- **Sarkis, M.**, Fyfe, A., Shah, N. & Papathanasiou, M.M. Towards 3-fold sustainability in biopharmaceutical supply chain planning: single-use vs stainless steel manufacturing. *To be submitted to ACS Sustainable Engineering*, In preparation.
- **Sarkis, M.**, Sachio, S., Kontoravdi, C., Shah, N. & Papathanasiou, M.M. (2024). Towards 3-fold sustainability in biopharmaceutical process development and product distribution, *Systems and Control Transactions*, 3, 75-83.
- **Sarkis, M.**, Fyfe, A., Kontoravdi, C. and Papathanasiou, M.M. (2024). Towards a Net Zero, socially sustainable and eco-efficient biopharma: how far are we?, *Current Opinion in Chemical Engineering* 44, 101027.
- **Sarkis, M.**, Bernardi, A., Shah, N. & Papathanasiou, M.M. (2022). Integrating environmental sustainability in next-generation biopharmaceutical supply chains, *Book of Abstracts of the 33rd European Symposium on Computer Aided Process Engineering*, 1633-1638.

7.1 Introduction

Capacity planning pushes towards manufacturing strategies that decrease therapy costs and maximize product availability. To this end, the previous Chapters 3-6 have focused on the development of frameworks that assist pharmaceutical planning under uncertainty to develop cost-effective and foster *uncertainty-awareness*.

Environmental sustainability awareness in supply chain for life sciences and pharmaceuticals is becoming increasingly relevant. Recent governmental and industrial initiatives to comply with Net Zero targets by 2030 (HM Government, 2021; AstraZeneca, 2022; Pfitzer, 2022). Many companies have already tackled some of the issues through strategies to reduce transport emissions throughout distribution, given that they represent 50-70% of emissions from the healthcare sector Pharmaceutical Technology (2023). This is, however, not a trivial task if the end-to-end supply chain is considered, in fact key challenges emerge at the manufacturing level too.

Many existing facilities use stainless steel processes and have lifetimes extending up to 20 years. Processes stainless steel multi-use (MU) equipment are water and energy intensive and require extensive use of solvents for cleaning and buffers in purification and formulation steps in downstream (DSP) and fill and finish (F&F) process sections. This results in significant waste water streams. Recent industrial trends propose a shift towards plastics-based single-use (SU) equipment to minimise upfront capital investment, help meet global demands faster with more adaptable modular equipment that increases flexibility. In this context, the use of disposable plastics has caused significant concern with regards to environmental sustainability, due to a high degree of solid contaminated waste (Samaras et al., 2022), as well as additional supply chain-related transport emissions related to the need to continuously source consumable equipment.

As a result, an emerging question in industry is whether SU equipment is more or less sustainable than MU manufacturing when accounting the manufacturing phase of the product of interest and the footprints related to the supply of key raw materials. Some studies have focused on quantifying environmental sustainability of these manufacturing alternatives (Budzinski et al., 2022; Pietrzykowski et al., 2013) with a focus on mAbs. Specifically, Life Cycle Assessment (LCA) studies have highlighted that in monoclonal antibodies (mAbs) manufacturing

platforms SU use reduces environmental footprint in terms of climate change, human health, ecosystem quality, natural resources and freshwater withdrawal compared to stainless steel (Pietrzykowski et al., 2013). This is predominantly relating to the removed requirement for clean- and steam-in-place (CIP/SIP), subsequently reducing use of energy-intensive water for injection (WFI) and resin utilisation. Specifically, the use of SU equipment leads to reductions up to 75% and 95% for WFI and purified water compared with stainless steel facilities. Additionally, SU introduction in DSP removes the need for caustic chemicals for post-processing treatment and sanitisation, which reduces the number of effluents in waste streams.

Beyond the quantification and comparison of manufacturing platforms, there is an opportunity to integrate these findings and leverage on LCA methods to then develop optimisation-based tools for *sustainability-aware* capacity plans for a range of manufacturing networks. These tools can help minimise footprints of manufacturing set ups of interest and quantify potential trade-offs with economic metrics. The review of literature contributions in the field presented in Chapter 2 effectively highlights that the problem of capacity planning and production planning has mainly been addressed for an optimisation of investment under process and demand uncertainty with no quantification of sustainability metrics. Key challenges tackled by works in this context (Gatica et al., 2003a; Tsang et al., 2007b; Lakhdar et al., 2005; Colvin and Maravelias, 2008) focus on clinical trial and demand uncertainty. In this sense, these approaches mainly focus on a minimisation of cost (*economic* sustainability) and maximisation of therapy availability (*social* sustainability).

In this chapter, we address this gap through the integration of LCA-based sustainability quantification a range of processes in a framework for the optimisation of economic metrics and environmental sustainability of manufacturing networks. Specifically, we introduce equipment selection (MU vs SU) as a decision variable within the optimisation to explore the performance of a range of manufacturing configurations. To conclude the analysis, we investigate the interconnectivity among cost-effective capacity planning, improved resource use and environmental sustainability.

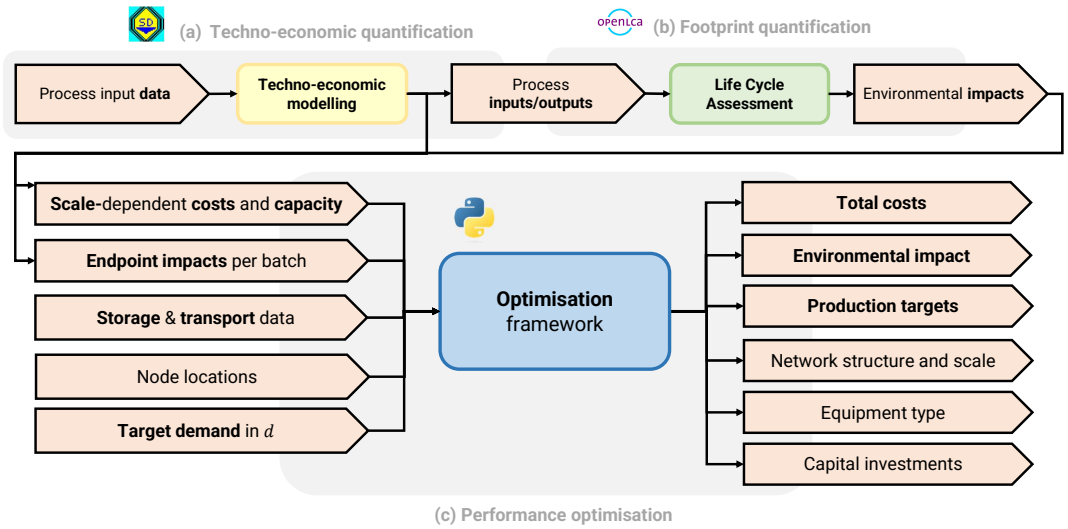


Figure 7.1: Quantification-optimisation framework: (a) techno-economic quantification (b) environmental footprint quantification, (c) performance optimisation.

7.2 Methodology

The framework presented in this Chapter consists of (i) a techno-economic performance quantification step, (ii) environmental footprint quantification and a (iii) optimisation step (Figure 7.1). The following sections illustrate the methodology adopted in the development of the quantification-optimisation framework and discuss the case studies used to illustrate the capabilities of the developed tool.

7.2.1 Problem statement

The problem considered in this study is focused on capacity planning and based on the network superstructure illustrated in Figure 7.2. The objective of the planning problem is to minimise total supply chain capital and operating costs and determine the following decisions:

- a network structure and links between nodes;
- investment in manufacturing and storage facilities;

- scales of manufacturing and number of parallel lines;
- production levels at each manufacturing node;
- transportation flows between existing nodes;
- supplied quantities at each demand node.

To do this, the optimisation considers the following:

1. a set T of time periods t;
2. a set N of nodes n;
3. a set A of manufacturing scales a
4. a set S of manufacturing process sections s;
5. a set L^S of manufacturing process lines l;
6. a set EQ of equipment types e;
7. a set MP of midpoint impacts w;
8. a set EP of endpoint impacts f;
9. datasets of scale-dependent batch sizes, batch times and costs;
10. datasets of scale-dependent midpoint and endpoint impacts per batch;
11. data on distances between each node n;
12. data on transportation costs;
13. data on storage costs;
14. demand in demand zones $n \in N^Z$ at time t.

The underlying challenge in this capacity planning problem is the potential trade-off that may emerge between cost, environmental sustainability and therapy availability. In this chapter, the trade-off amongst the three is explored.

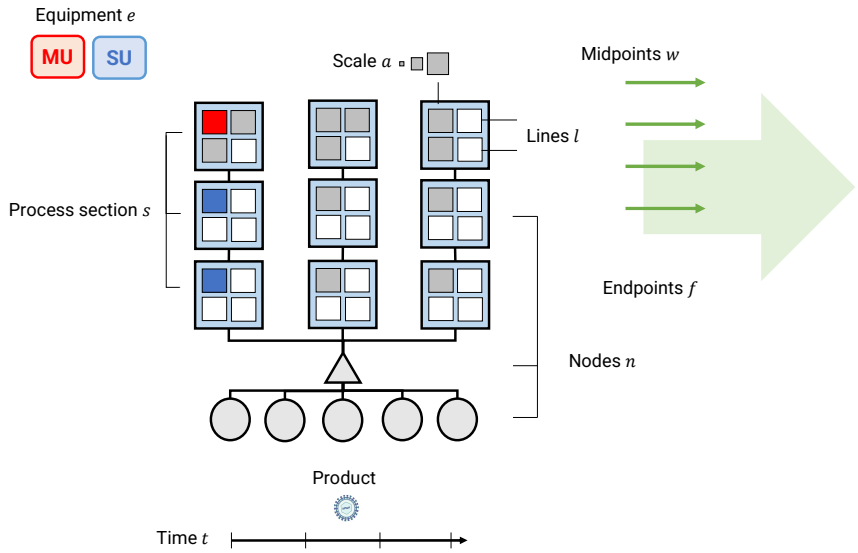


Figure 7.2: Superstructure for optimisation of LV-based case study. Nodes n comprise manufacturing nodes (square), storage (triangle) and demand zones (circles). Each manufacturing node is modelled as a set of process sections s comprising parallel manufacturing lines l at scale a, implementing equipment type e and varying by time period t.

7.2.2 Environmental footprint quantification

The quantification of environmental footprint is done through a Life Cycle Assessment (LCA) and is implemented in OpenLCA. The developed process models use the Life Cycle Inventory (LCI) database EcoInvent 3.6 APOS and the impact assessment method ReCiPe 2016. Details regarding the process models are provided in Appendix D.

Through this method, both midpoint and endpoints are considered as environmental key performance indicators (KPIs). In this chapter, the focus is on manufacturing platforms alternatives and environmental footprints can be lowered through the selection of suitable equipment. A process was created for each scale a, equipment e and process section s within manufacturing nodes $n \in N^J$. In this fashion, the LCA helps quantify the endpoint impacts e^{EP}_{asef} and midpoint impacts e^{MP}_{asem} per batch. These parameters can be integrated in constraints which computing environmental footprints for candidate network structures within the optimisation formulation.

7.2.3 Optimisation formulation

The optimisation step in Figure 7.1 considers a multi-period single-product planning problem where $T = \{T_1, T_2, T_3, T_4, T_5\}$ the considered product is a lentivirus-based product for gene therapy applications. In this chapter we consider a deterministic formulation, thereby introducing a single scenario realisation for the uncertain parameters and for demand scenarios. An additional complexity is added by considering possibility to install different equipment $EQ = \{SU, MU\}$.

7.2.3.1 Objective function

The objective of the optimisation is to minimise the total cost z of the supply chain, which in this Chapter is a deterministic quantity (Eq.7.1). Total costs include capital costs and operating costs of manufacturing, storage costs and transport costs.

$$\min \; z = TC^{CAPj} + TC^{OPj} + TC^{OPk} + TC^{T} \tag{7.1}$$

Where TC^{CAPj} is the total capital manufacturing cost, TC^{OPj} is the total operating manufacturing cost, TC^{OPk} is the total operating storage cost and TC^{T} is the total transport.

7.2.3.2 Capital manufacturing costs

Capital costs of manufacturing (Eq.7.2) are computed as a sum of facility installation costs and product-specific equipment investment costs.

$$TC^{CAPj} = \sum_{t \in T} \sum_{n \in N^J} C_{tn}^{EQj} \tag{7.2}$$

The term in Eq. 7.2 is C^{EQj}, which is the total cost of equipment investment at time t, in node $n \in N^J$.

$$C_{tn}^{EQj} = \sum_{e \in EQ} \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} c_{as}^{EQjl} H_{tnalse} \quad \forall t, n \in N^J \tag{7.3}$$

The equipment installation cost C^{EQj} is computed as a function of binaries H, which mark whether equipment for the manufacturing of product at scale a is installed in lines l of equipment e of section s in plant $n \in N^J$ at time t and capital cost parameters c^{EQjl} (Eq.7.3):

7.2.3.3 Operating manufacturing costs

Total operating costs TC^{OPj} (Eq. 7.4) of manufacturing are calculated as the sum of annual variable costs C^{VARj}, facility related costs C^{FACj} and labour costs C^{LABj} for each manufacturing plant $n \in N^J$ at time t.

$$TC^{OPj} = \sum_{t \in T} \sum_{n \in N^J} C_{tn}^{LABj} + C_{tn}^{VARj} + C_{tn}^{FACj} \tag{7.4}$$

Labour costs C^{LABj} for each time period t, manufacturing plant $n \in N^J$ and are computed (Eq. 7.5) as a summation of the (i) labour costs c^{LABjl} parameters for each scale a, process section l and equipment e multiplied by (ii) batches B respectively which vary with respect to time t, plant $n \in N^J$, scale a, equipment e, process section s and production line l:

$$C_{tn}^{LABj} = \sum_{e \in EQ} \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} c_{ase}^{LABjl} B_{tnasle} \quad \forall t, n \in N^J \tag{7.5}$$

Variable costs C^{VARj} for each time period t and manufacturing plant $n \in N^J$ are computed (Eq. 7.6) as a multiplication of the (i) variable costs parameters c^{VARjl} for each scale a, equipment e and process section s by (ii) batches B respectively which vary with respect to time t, plant $n \in N^J$, scale a, process section s, production line l and equipment e:

$$C_{tn}^{VARj} = \sum_{e \in EQ} \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} c_{ase}^{VARjl} B_{tnasle} \quad \forall t, n \in N^J \tag{7.6}$$

Facility-dependent costs C^{FACj} are dependent on depreciation and equipment maintenance costs. In Eq. 7.7, facility-dependent costs for every manufacturing plant $n \in N^J$ and time t are computed as multiplication of (i) facility costs parameters c^{FACjl} for each scale a, process section l and equipment e by (ii) equipment installation-related binaries M respectively which vary with respect to time t, plant j, scale a, process section s, production line l and equipment e.

$$C_{tn}^{FACj} = \sum_{e \in EQ} \sum_{a \in A} \sum_{s \in S} \sum_{l \in L^S} c_{ase}^{FACjl} M_{tnalse} \quad \forall t, n \in N^J \tag{7.7}$$

7.2.3.4 Logistics-related costs

Total operating costs of storage TC^{OPk} are a function of (i) the batches handled in the process Q, which corresponds to the number of batches of scale a of product shipped from upstream storage node $n \in N^K$ to downstream demand zone $m \in N^Z$, (ii) the cost parameter c^{OPk} at scale a.

$$TC^{OPk} = \sum_{t \in T} \sum_{n \in N^K} \sum_{m \in N^Z} \sum_{a \in A} c_a^{OPk} Q_{tnma} \tag{7.8}$$

Total transport related costs TC^T are calculated as a function of (i) the distance traveled ρ_{nm} between upstream nodes n and downstream nodes $m \in M^N$, (ii) the number of batches Q at scale a shipped from n to $m \in M^N$, (iii) a cost parameter for transport c^T.

$$TC^T = c^T \sum_{t \in T} \sum_{n \in N^M} \sum_{m \in N^M} \sum_{a \in A} \rho_{nm} Q_{tnma} \tag{7.9}$$

7.2.3.5 Logic constraints

Logic constraints govern the construction of feasible supply chain network structures. First, a link between an upstream node n and downstream node $m \in M^N$ is established only if the upstream node n exists at time t.

$$X_{tnm} \leq Y_{tn} \quad \forall t, n \in N, m \in M^N \tag{7.10}$$

Where X is the binary variable that indicates the establishment of a link between nodes n to $\in M^N$ and Y is the binary variable that indicates the existence of a node in location $n \in N$ at time t.

Sourcing constraints ensure that each node is sourced by at least one route via the binary variables X. In the case of storage nodes $m \in N^K$, the node exists if the binary variable $Y = 1$ (Eq. 7.11). In the case of demand zones $m \in N^Z$ the nodes are sourced at all times, with the R.H.S. of Eq.7.12 equal to 1.

$$\sum_{n \in N^J} X_{tnm} \geq Y_{tm} \quad \forall t, m \in N^K \tag{7.11}$$

$$\sum_{n \in N^K} X_{tnm} \geq 1 \quad \forall t, m \in N^Z \tag{7.12}$$

In this problem, we consider that at most one scale per plant $n \in N^J$ can be installed. Hence, the following constraint enforces that a scale a is selected only if the manufacturing node exists. Furthermore it enforces the selection of only 1 scale of manufacturing, since the L.H.S. summation cannot take values larger than 1:

$$\sum_{a \in A} V_{tna} \leq Y_{tn} \quad \forall t \in T, n \in N^J \tag{7.13}$$

Where V is the binary variable that indicates the selection of scale a in location $n \in N^J$ at time t.

Similarly, a production line l made of equipment e in process section s at scale a is established only if the corresponding scale is established at that location $n \in N^J$ and time t:

$$Z_{tnasle} \leq V_{tna} \quad \forall t, n \in N^J, a, s, l, e \tag{7.14}$$

Where Z takes a value of 1 if a line l of equipment e is established in process section s, at scale a and time t. Equipment installation can occur only if the manufacturing line is available and installed via the binary variable Z. Therefore:

$$M_{tnasle} \leq Z_{tnasle} \quad \forall t, n \in N^J, a, s, l, e \tag{7.15}$$

Where M indicates the product-specific equipment installation at time t, in node $n \in N^J$, at scale a, process section s and line l of equipment e.

In a similar fashion, product allocation W on the L.H.S. of Eq. 7.16 can take value of 1 only if the production line is installed, which is modelled via $M = 1$.

$$W_{tnasle} \leq M_{tnasle} \quad \forall t, n \in N^J, a, s, l, e \tag{7.16}$$

Finally, for each manufacturing line l in each production node $n \in N^j$, scale a, time t only one type of equipment can be installed. This modelled via the following equation:

$$\sum_{e \in EQ} Z_{tnasle} \leq 1 \quad \forall t, n \in N^J, a, s, l \tag{7.17}$$

7.2.3.6 Multi-period scale-up constraints

The scale-up decisions are constrained in the case of MU. Once a facility is installed the equipment scale will remain fixed for the subsequent time periods. To model this the following constraint is written:

$$Z_{tnasle} = Z_{t-1,nasle} + H_{tnasle} \quad \forall t|t > 1, n \in N^J, a, s, l, e|e = MU \tag{7.18}$$

Where the R.H.S. corresponds to the existence of a line l in process section s in plant j and scale a at time $t > 1$. This variable takes a value of 1 if it was activated in the previous time period (L.H.S.) or invested in the current t via the investment variable A and takes a value of 0 otherwise.

Conversely, this does not hold for SU based equipment, where equipment of larger scale can be flexibly installed/de-installed from one period t to the next. Hence we relax Eq. 7.18 as follows:

$$Z_{tnasle} \leq Z_{t-1,nasle} + H_{tnasle} \quad \forall t|t > 1, n \in N^J, a, s, l, e|e = SU \tag{7.19}$$

7.2.3.7 Transport constraints

Transport constraints regulate the maximum and minimum achievable flows between upstream nodes n and downstream nodes $m \in M^N$ in the supply chain network at time t.

$$Q^{MIN} X_{tnm} \leq Q_{tnma} \leq Q^{MAX} X_{tnm} \qquad \forall t, n \in N, m \in M^N \tag{7.20}$$

Where Q^{MIN} and Q^{MAX} correspond to lower and upper bound parameters for the flow variables Q if the link is established via $X = 1$. Instead, these are enforced to be equal to 0 if $X = 0$.

7.2.3.8 Manufacturing constraints

Time constraints bound the time allocated to production in each process section s, manufacturing line l, process plant $n \in N^J$ at all times t.

$$T^{MIN} W_{tnasle} \leq T_{tnasle} \leq T^{MAX} W_{tnasle} \quad \forall t, j, a, l, s, e \tag{7.21}$$

Where the bounds are activated if the product allocation variable $W = 1$, otherwise the variable T is forced to be equal to 0.

The number of batches B processed in each process section s, process line l, equipment e at scale a in location $n \in N^J$ time t are calculated as a function of (i) the time allocation variable T, (ii) process parameter realisations. Specifically, these correspond to the process time taken to collect the first batch in the campaign α in each process section s and the length of the bottlenecks within each process section τ.

$$B_{tnasle} = W_{tnasle} + 1/\tau_{sp}(T_{tnasle} - \alpha_{sp}W_{tnasle}) \quad \forall t, n \in N^J, a, s, e \tag{7.22}$$

The above expression (Eq. 7.22) is written to compute the number of batches B in each campaign as a sum of the first batch ($W = 1$) and the remaining $B-1$ batches which are obtained within the time slot $T - \alpha$.

7.2.3.9 Network constraints

The overall material balance is written for every node in the supply chain network as follows:

$$I_{tna} = I_{tna|t-1} + Q^{IN}_{tna} - Q^{OUT}_{tna} + P_{tna} - S_{tna} \quad \forall t, n \in N, a \tag{7.23}$$

Where I is the inventory of batches at scale a held at time t, in node n, Q^{IN} and Q^{OUT} denote the inflow of batches from upstream nodes and outflow to downstream nodes, P denotes the amount of material produced and S denotes the amount of material consumed (i.e. sold). This general form of material balance can be re-written for each node type. In this Chapter, we consider no inventory accumulation hence $I = 0$.

In the case of manufacturing facilities $n \in N^J$, only production and outbound flow occur; therefore, the material balance can be written as:

$$P_{tna} = Q^{OUT}_{tna} \quad \forall t, n \in N^J, a \tag{7.24}$$

Where the production term P corresponds of a summation of the batches processed in the F&F process section through lines l, at scale a, in plant $n \in N^J$ at time t.

$$P_{tna} = \sum_{l \in L^F} \sum_{e \in EQ} B_{tjasle|s=F} \quad \forall t, n \in N^J, a \tag{7.25}$$

The outbound flow for manufacturing corresponds to a summation of Q from nodes $n \in N^J$ to storage nodes $n \in N^K$.

$$Q^{OUT}_{tna} = \sum_{m \in N^K} Q_{tnma} \quad \forall t, n \in N^J, a \tag{7.26}$$

Furthermore within a manufacturing plant a set of additional mass balances is written to regulate the flow of batches between process sections s. Hence:

$$J_{tnas} = J_{tnas|t-1} + \sum_{l \in L^S} \sum_{e \in EQ} B_{tnasle} - \sum_{l \in L^S} \sum_{e \in EQ} B_{tna,s+1,le} \quad \forall t, n \in N^J, a, s, e \tag{7.27}$$

Where J corresponds to the inventory of batches held between process section s upstream and its downstream process section $s+1$. This holds for upstream to downstream process section and downstream to F&F parts of the process. No inventory is considered, hence $J = 0$.

In the case of storage facilities $n \in N^K$, material is not produced or consumed and the expression in Eq. 7.23 simplifies to a balance between inbound and outbound flows.

$$Q^{IN}_{tna} = Q^{OUT}_{tna} \quad \forall t, n \in N^K, a \tag{7.28}$$

Where the inbound flows Q^{IN} of storage correspond to a summation of Q from nodes $n \in N^J$ to the storage node $n \in N^K$:

$$Q^{IN}_{tma} = \sum_{n \in N^J} Q_{tnma} \quad \forall t, m \in N^K, a \tag{7.29}$$

Conversely, the outbound flows Q^{OUT} of storage correspond to a summation of Q the node $n \in N^K$ to demand zones $n \in N^Z$.

$$Q^{OUT}_{tna} = \sum_{m \in N^Z} Q_{tnma} \quad \forall t, n \in N^K, a \tag{7.30}$$

Finally, in the case of demand zones $n \in N^Z$, the overall material balance in Eq. 7.23 simplified to a balance between inbound flows and product sales as follows:

$$Q^{IN}_{tna} = S_{tna} \quad \forall t, n \in N^Z, a \tag{7.31}$$

Where Q^{IN} is computed as a summation of Q from upstream storage nodes $m \in N^K$ to each demand zone $n \in N^Z$:

$$Q^{IN}_{tma} = \sum_{n \in N^K} Q_{tnma} \quad \forall t, m \in N^Z, a \tag{7.32}$$

The amount of doses supplied to each demand zone corresponds to the amount of batches S at demand zone $n \in N^Z$ at time period t produced with scale a multiplied by the batch size parameter χ. Doses delivered are constrained to be greater than or equal that the demand D on the R.H.S at the demand zone $n \in N^Z$.

$$\sum_{a \in A} S_{tna}\chi_a \geq D_{tn} \quad \forall t, n \in N^Z \tag{7.33}$$

In this study, the demand D is calculated by the following equation based on the total demand for the product of interest D^T:

$$D_{tn} = D^T_t / \text{card}(N^Z) \quad \forall t, n \in N^Z \tag{7.34}$$

7.2.3.10 Environmental constraints

The footprint of processing a batch in section s of a process of scale a and equipment type e can be computed both at midpoint and endpoint level through the following constraints.

The impact for each midpoint $w \in MP$ can be computed as the multiplication of the batches manufactured in each process equipment e, process line l, scale a, process section s and plant $n \in N^j$ and the quantified impacts e^{MP} through the LCA midpoint impact quantification step.

$$E^{MP}_{tw} = \sum_{e \in EQ} \sum_{n \in N^j} \sum_{s \in S} \sum_{l \in L} e^{MP}_{asew} B_{tnasle} \quad \forall w, t \tag{7.35}$$

Similarly the impact for each endpoint $f \in EP$ can be computed as the multiplication of the batches manufactured in each process equipment e, process line l, scale a, process section s and plant $n \in N^j$ and the quantified impacts e^{EP} through the LCA endpoint impact quantification step.

$$E_{tf}^{EP} = \sum_{e \in EQ} \sum_{n \in N^j} \sum_{s \in S} \sum_{l \in L} e_{asef}^{EP} B_{tnasle} \quad \forall f, t \tag{7.36}$$

Finally the total environmental footprint TE is computed within the optimisation after normalisation by the monetisation factors (Appendix D):

$$TE = \sum_{t \in T} \sum_{f \in EP} \mu_f E_{tf}^{EP} \tag{7.37}$$

Given the logic constraints in Eq. 7.14, 7.15, 7.16 and 7.17, production is constrained to happen only for a single scale, a single process equipment type in each production equipment in process section s, manufacturing plant $n \in N^j$, line l and time t. Performing an economic minimisation corresponds to selecting the required capacity and and equipment type, which minimise cost and evaluates a sustainability metric for each solution through the environmental constraints. Conversely, a single-objective optimisation with respect to environmental footprint can also be performed, whereby footprints are reduced and capacity and equipment is selected to meet demand D_t^T. Effectively, environmental footprints of interest correspond to midpoints E^{MP} (e.g. global warming, water consumption) or normalised quantities including endpoints E^{EP} and monetised TE values.

7.2.4 Computational resources

The framework is built in Python 3.11 implementing the PYOMO package. The MILP is solved with GUROBI v11 on a Intel(R) Core(TM) i7-10700 CPU @ 2.90GHz machine with 64.0 GB RAM. Model statistics are displayed in Table 7.1.

Table 7.1: Computational statistics of the snapshot optimisation framework

Model	**CPU (s)**	**Constraints**	**Integer (Binary)**	**Continuous**	**Gap**
MU	3s	16734	6070(4630)	5359	0%
SU	2s	16734	6070(4630)	5359	0%

7.2.5 Uncertainty quantification

Each of the LCA flowsheets developed requires user-inputs on the key input and output flows (Appendix D), which define the foreground system of the process and uses extensive background databases to calculate background inventories. Foreground inventories are inherently uncertain and Chapter 3 extensively discusses the magnitude of scale-dependent process uncertainty for the emerging products studied in this Thesis. In addition to this, the models proposed are generalised and it is important to account for variability amongst companies in the material quantities used (e.g. equipment weights for SU and MU manufacturing). To address, this we consider triangular distributions for all foreground inventories, whereby upper bound and lower bound correspond to ±50% from the central nominal value. Background inventories carry uncertainty which is modelled through log-normal distributions in OpenLCA.

We perform a total of 1000 Monte Carlo (MC) runs for each of the processes of interest to quantify midpoint and endpoint uncertainty. Each run assumes that processes of interest are fully independent, which coincides with the assumptions underpinning the uncertainty analyses presented in Chapter 3. In these contexts, independent sampling translates into uncertainty quantification that disregards potential correlations between the two or more processes, allowing an initial understanding process-specific environmental footprints and the variability of each of them. This assumption of *independence* may break down in LCA studies, as processes effectively share the same resources uncertainty (i.e. background suppliers and inventories). Quantifying inter-process correlation would require the implementation of dependent sampling, translating into a single LCA model comprising all processes and implementing an integrated MC simulation.

7.3 Results & discussion

7.3.1 Quantification of environmental footprint

The endpoint quantification obtained with LCA for SU and MU is displayed in Figure 7.3, showing contributions by midpoint categories and process sections. Furthermore, inventory flow contributions are displayed in Figures 7.4 and 7.5.

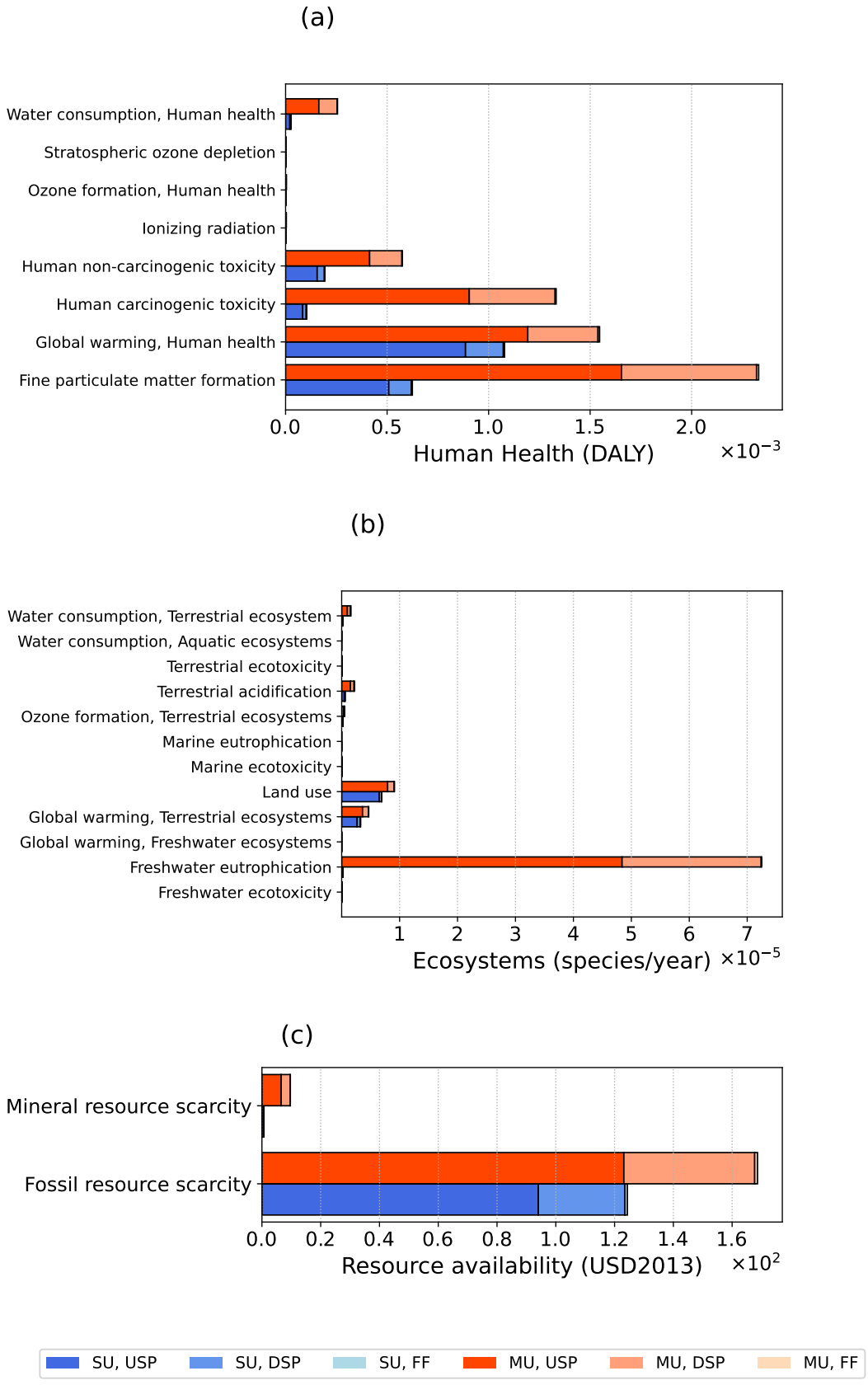


Figure 7.3: Process sections contributions to SU and MU processes at scale 2000L on endpoints of (a) human health, (b) ecosystems and (c) resource availability.

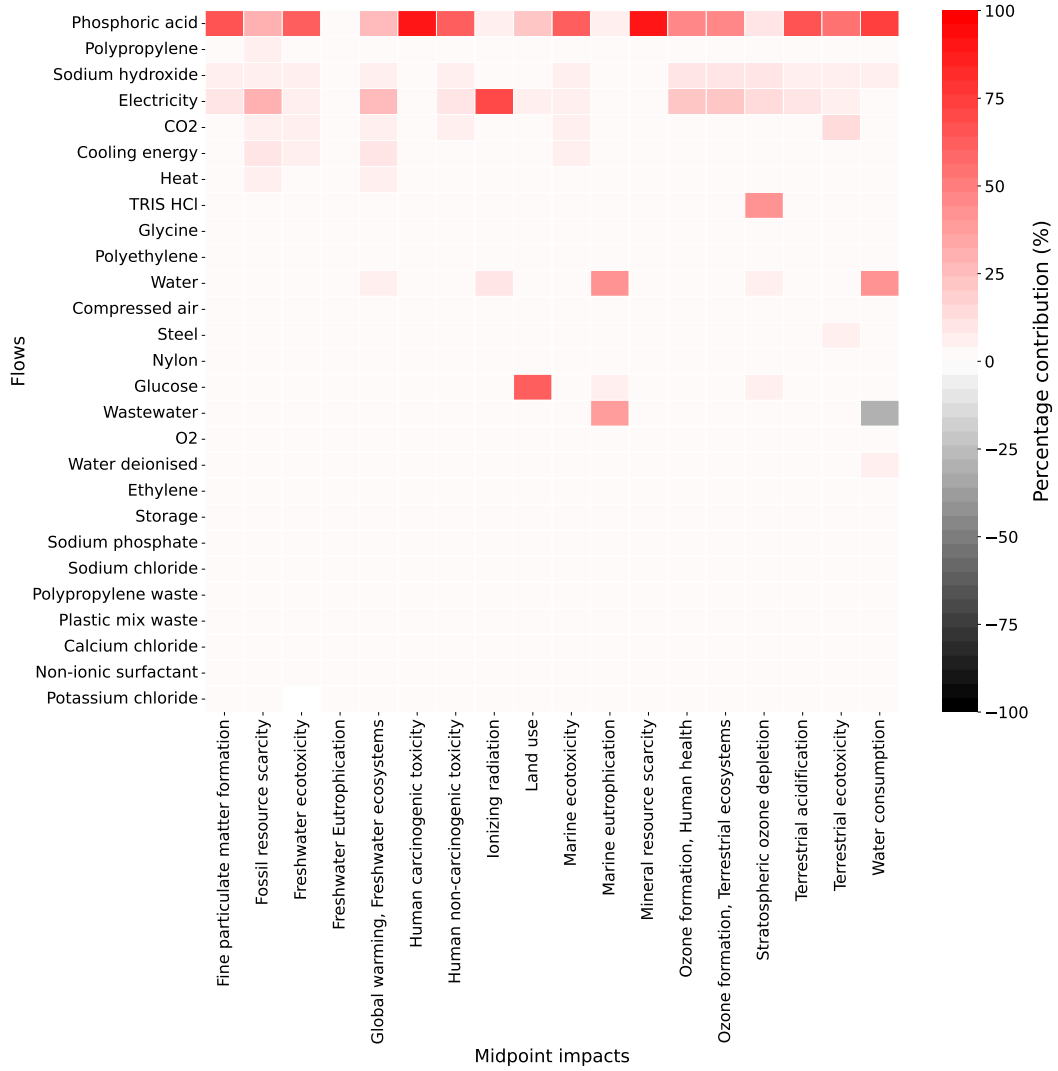


Figure 7.4: Inventory flows percentage contributions in MU process at scale 2000L to midpoints.

Overall, the MU process is more impactful compared to the SU process. This holds for all endpoint impact indicators including human health (Figure 7.3a), ecosystems (Figure 7.3b) and resource availability (Figure 7.3c). The impacts are a result of the use of electricity and phosphoric acid in SIP/CIP steps (Figure 7.4), which largely impact most midpoint categories with percentages in the range of 25-100%. Interestingly, in SU facilities the larger impact contributors are not linked to the plastics used as consumables. As can be seen in Figure 7.5, the electricity used to run the facility results in the largest percentage contribution across all impact categories, with shares of 54-100% compared to values below 15% for plastics-related flows. Effectively, through this quantification of critical sources of impact, manufacturers could look into reducing environmental

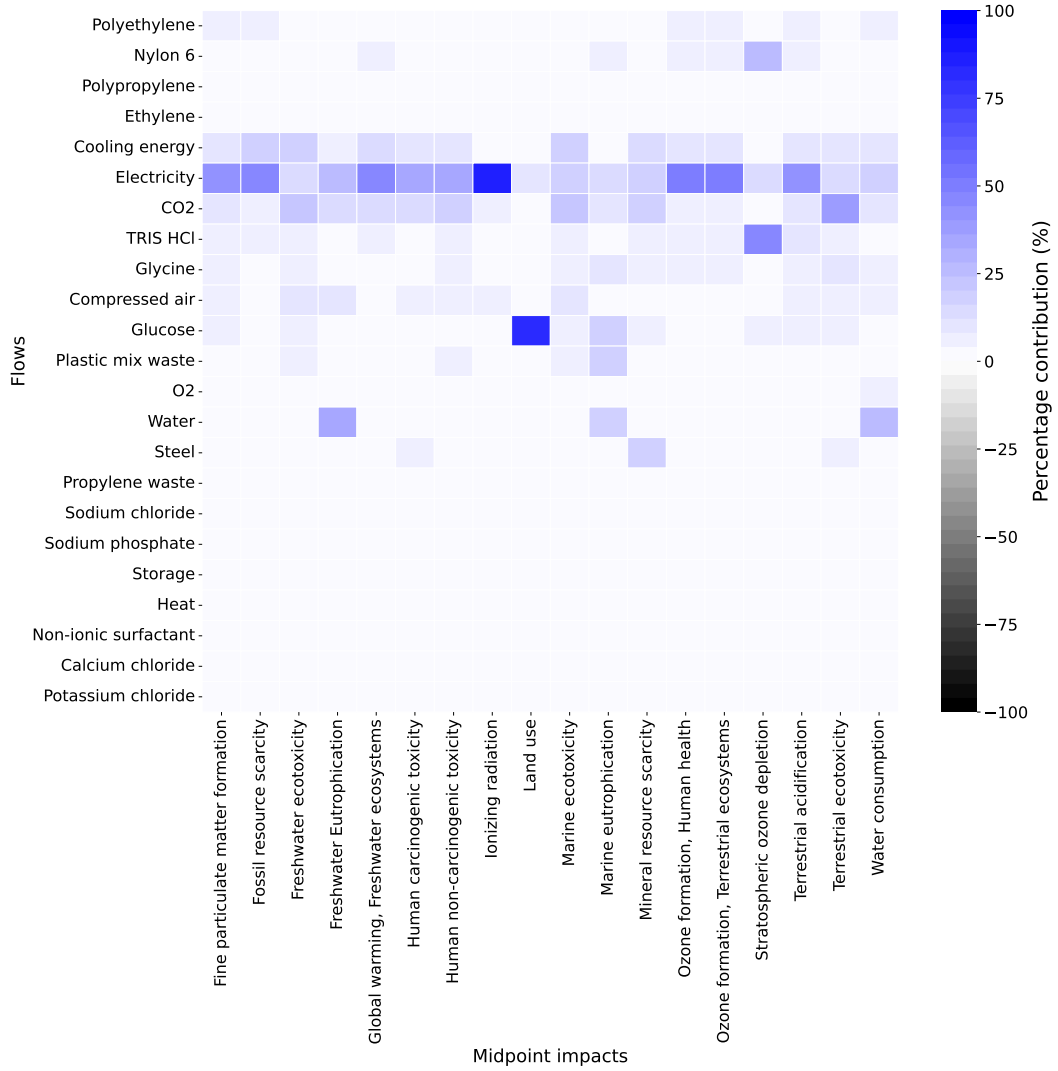


Figure 7.5: Inventory flows percentage contributions in SU process at scale 2000L to midpoints.

footprints of existing processes through process intensification and reducing the amount used per functional unit.

In Figure 7.3, the respective contribution of USP, DSP and F&F processes is reported. It emerges that the largest part of the footprint is due to USP activities. For the MU process this could be justified by the largest use of electricity to run stirred tank reactors in USP parts of the process and the phosphoric acid used for SIP/CIP. In comparison, DSP MU tanks are smaller due to smaller volumetric flows, thereby reducing the amount of WFI and acid solution used during cleaning. For the SU case, the prevalence of USP footprints is again justified by an increased use of electricity in fermentation steps. In comparison, the absence of phosphoric acid results in consistently lower footprints for SU. Of course, in

the instance of human health and resource availability (Figure 7.3a and c), the percentage footprint reduction is overall less significant compared to the endpoint impact on ecosystems (Figure 7.3b). Ultimately, the percentage difference in footprint between SU and MU would depend on the value of the normalisation factors μ_f which indicate the relative importance of endpoint indicators within the unified environmental footprint metric.

7.3.2 Deterministic optimisation results

Given the quantified environmental footprints for different process sections and scales, we present the results obtained by solving a multi-period deterministic constrained optimisation to compare networks implementing SU equipment as opposed to MU equipment. Effectively, the SU constrained solution corresponds to solution of the unconstrained cost optimisation, where equipment selection binaries are degrees of freedom. This is a result of SU being the most cost-effective option, as discussed in Chapter 6, as overall equipment-related capital and operating costs outweigh the increased costs of consumables.

As can be seen in Figure 7.6, the network capacity and parallel lines differ between the SU and MU case. In this deterministic multiperiod problem, future demands are assumed to be known for certain, therefore both solutions anticipate the installation capacity (Figure 7.6a and f). In the SU case, 2 200L lines supply the product in $t = T_1$, which is then stored in the subsequent time period (Figure 7.6b). Given the possibility to install and de-install capacity of different scales, subsequent time periods, use a new set of 3 manufacturing lines at 2000L (Figure 7.6c-d). Conversely, the MU investment constraints force the existence of manufacturing lines in subsequent time periods, after they are first installed. As can be seen in Figure 7.6f-l, additional capacity is installed cumulatively.

The economic and environmental performance of the two scale-up strategies is displayed in Figure 7.7b-h. As demand increases the total cost and total environmental footprint of the supply chain increases. MU has the largest total cost and largest total environmental footprint (Figure 7.7b,c,e and g), which result in larger unit manufacturing costs (Figure 7.7d) and unit emissions cost (Figure 7.7f and h). In both cases, an overall benefit of economies of scale is seen, with unit cost are significantly reduced for both MU and SU networks as soon as the demand allows for better utilisation of manufacturing resources. This trend

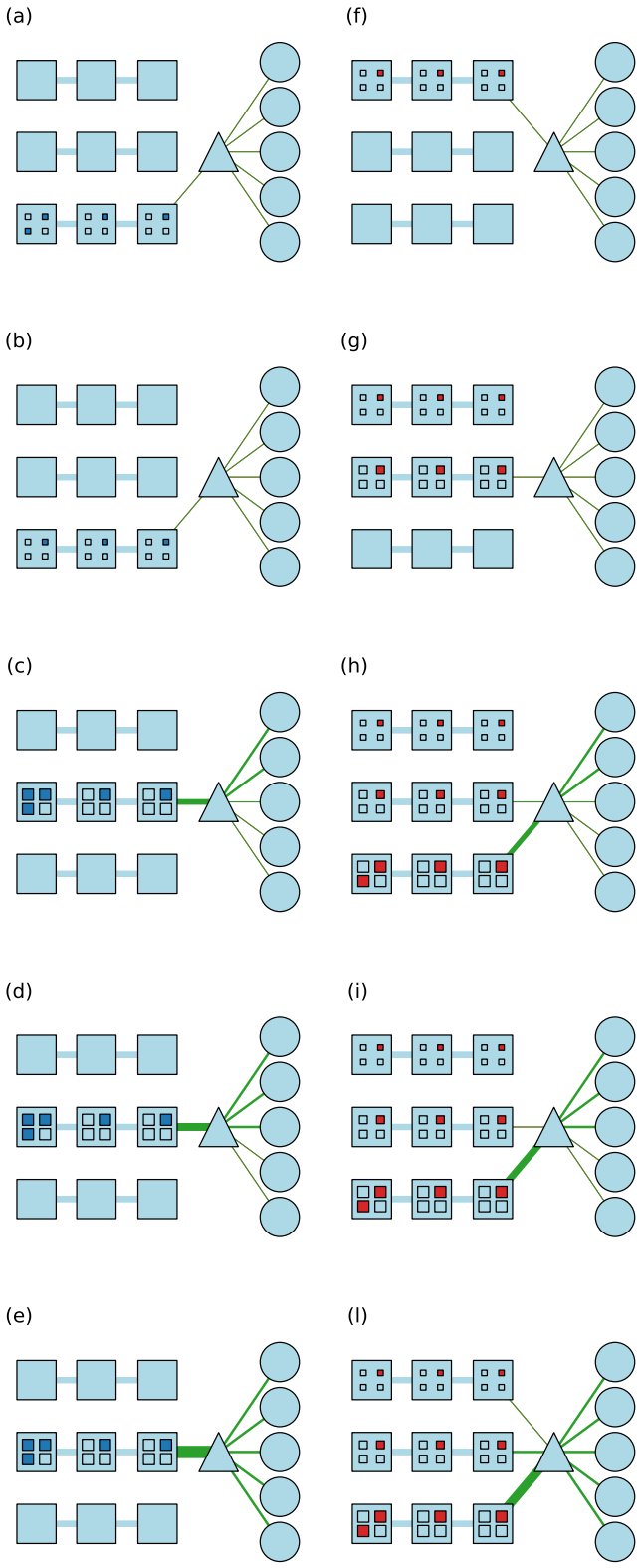


Figure 7.6: Scale-up structures for SU case (a) $t = T_1$, (b) $t = T_2$, (c) $t = T_3$, (d) $t = T_4$, (e) $t = T_5$ and MU case (f) $t = T_1$, (g) $t = T_2$, (h) $t = T_3$, (i) $t = T_4$, (l) $t = T_5$. MU equipment selected (■), SU equipment selected (■) and inactive lines (□)

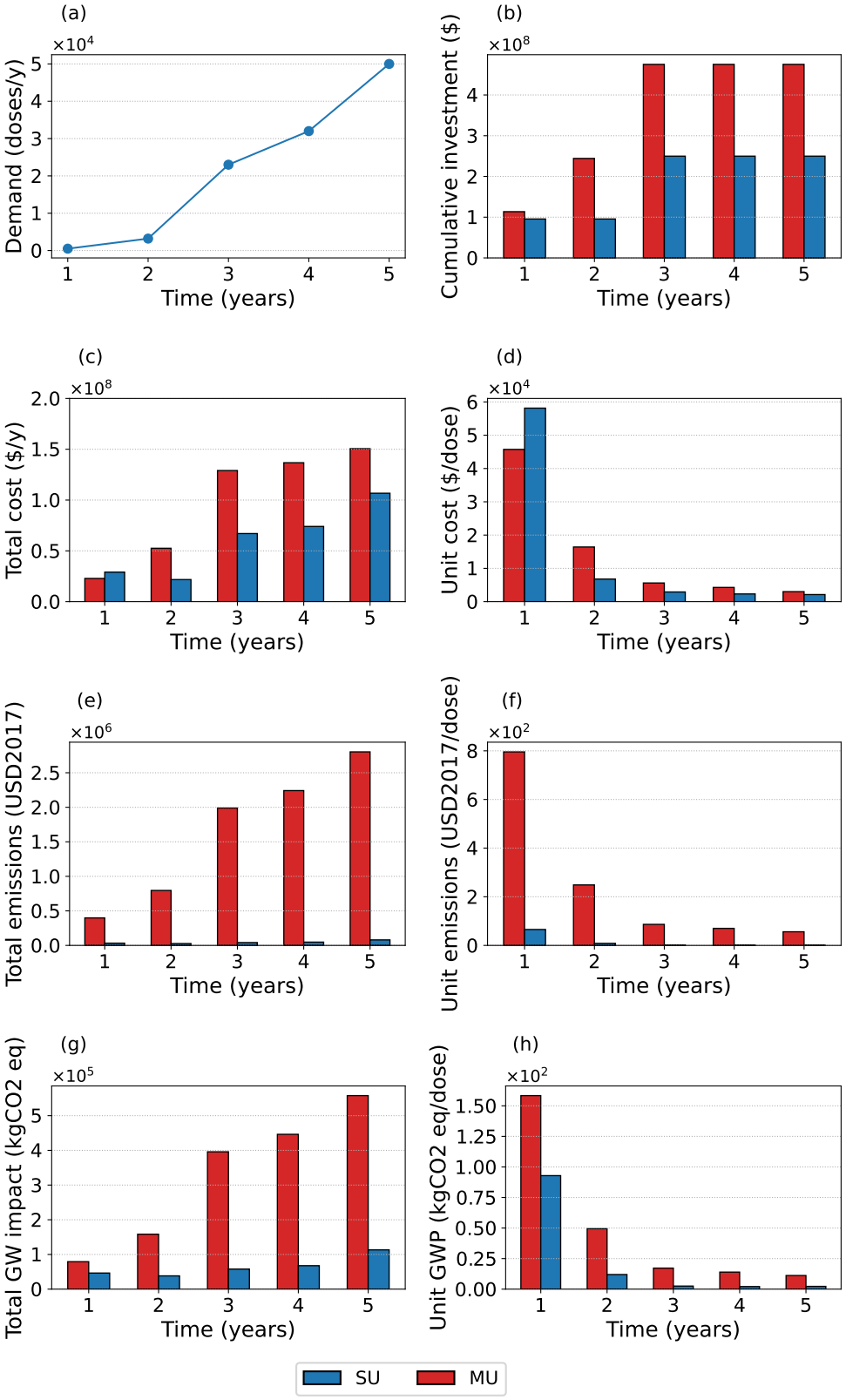


Figure 7.7: Scale up KPIs:(a) input demand, (b) cumulative capital investment, (c) total cost, (d) unit cost, (e) total footprint, (f) unit footprint, (g) total GWP emissions, (h) unit GWP emissions

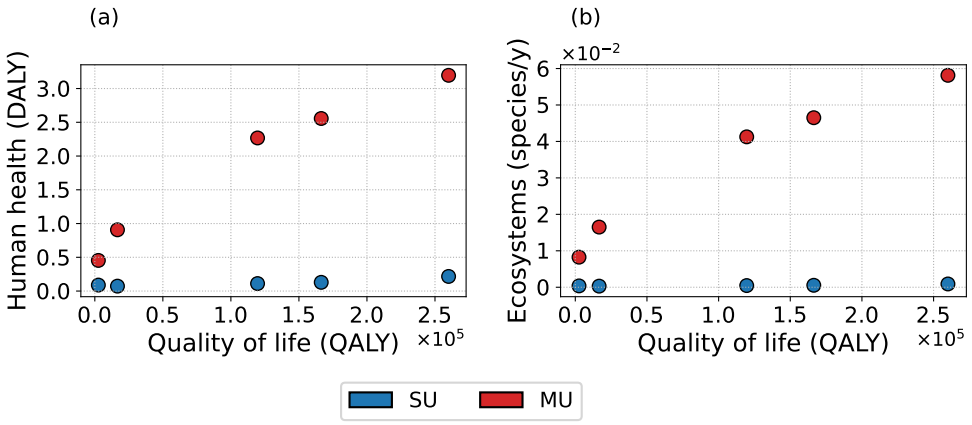


Figure 7.8: Social vs environmental benefit: (a) quality of life improved vs impact on human health, (b) quality of life improved vs impact on ecosystems.

is also seen with respect to the environmental indicators.

Another key output of the deterministic optimisation is the comparison between the benefit of meeting a demand to cure life-threatening diseases and having an impact on the environment. As argued, total costs and environmental footprints increase as demands increase. Figure 7.8a displays a trade-off curve between metrics of quantifying the lives saved through product deployment and the lives lost due the manufacturing activities. For each gene therapy delivered, a QALY corresponds to 5.2 years of additional life years (Wong et al., 2023). This means that by supplying 50,000 doses (t = T_5) the manufacturing network saves a cumulative amount of $> 2,500$ years of health. In doing so, it causes a damage on human health of 3 DALYs in the MU case, which correspond to 3 years lost of good health, which orders of magnitude lower than the benefit. In the SU case, this value is near 0. Similary, the number of species becoming extinct per year is orders of magnitude lower than the benefit to human health. This indicates that in either case, the optimisation model suggests that social benefits may outweigh environmental impacts concerning the impact on human life and animal species.

The challenge remains the minimisation of all other indicators, for instance linked to global warming, water pollution and resource availability. Overall, SU equipment is shown to be the most cost-effective and eco-efficient alternative. Nevertheless, if MU manufacturing is the only available option, efforts to avoid sub-optimal outcomes will result in lower emissions per product. On the

one hand, better resource utilisation results from economies of scale and process intensification, whereby the amount of resources used per unit product manufactured are reduced. On the other hand, improved resource utilisation is achieved through a careful assessment of projected demands and the impact of uncertainty on costs and environmental footprints. As shown in Chapter 4, deterministic optimisation disregarding uncertainty may lead to sub-optimal resource utilisation and failure to meet target demands. If footprints per dose are lowered through manufacturing planning aimed at maximising utilisation, the development of *uncertainty aware* capacity plans is crucial to lower environmental impact per dose wherever possible.

7.3.3 Quantification of footprints uncertainty

The integration of sustainability metrics in the problem of capacity planning under uncertainty introduces significant complexities. Effectively, uncertainties discussed so far in this thesis are cost-, throughput- and demand-related and are related mainly to cost and social objectives. The quantification of environmental footprints carries significant uncertainties tied to both calculated foreground and background inventories of the LCA method. Here, we investigate the variability of footprints of each process if foreground and background uncertainties are considered, sampling from the ranges displayed in Appendix D and recording resulting midpoint and endpoint model outputs.

Figure 7.9 presents a comparison between some representative environmental impact categories for the scales of interest comparing MU and SU respectively and display the quantified uncertainty. As mentioned in the methodology, the displayed uncertainty ranges were obtained through independent sampling of each process, treating each input parameter as statistically independent and meaning that changes in one parameter do not influence others. This approach could lead to results which do not provide an accurate representation of real-world systems and interactions, possibly underestimating or overestimating variability. Through this approach one can only quantify the uncertainty in environmental footprints for each independent processes and gain an initial understanding of footprints variability, whilst assuming that processes do not share common upstream providers and act independently.

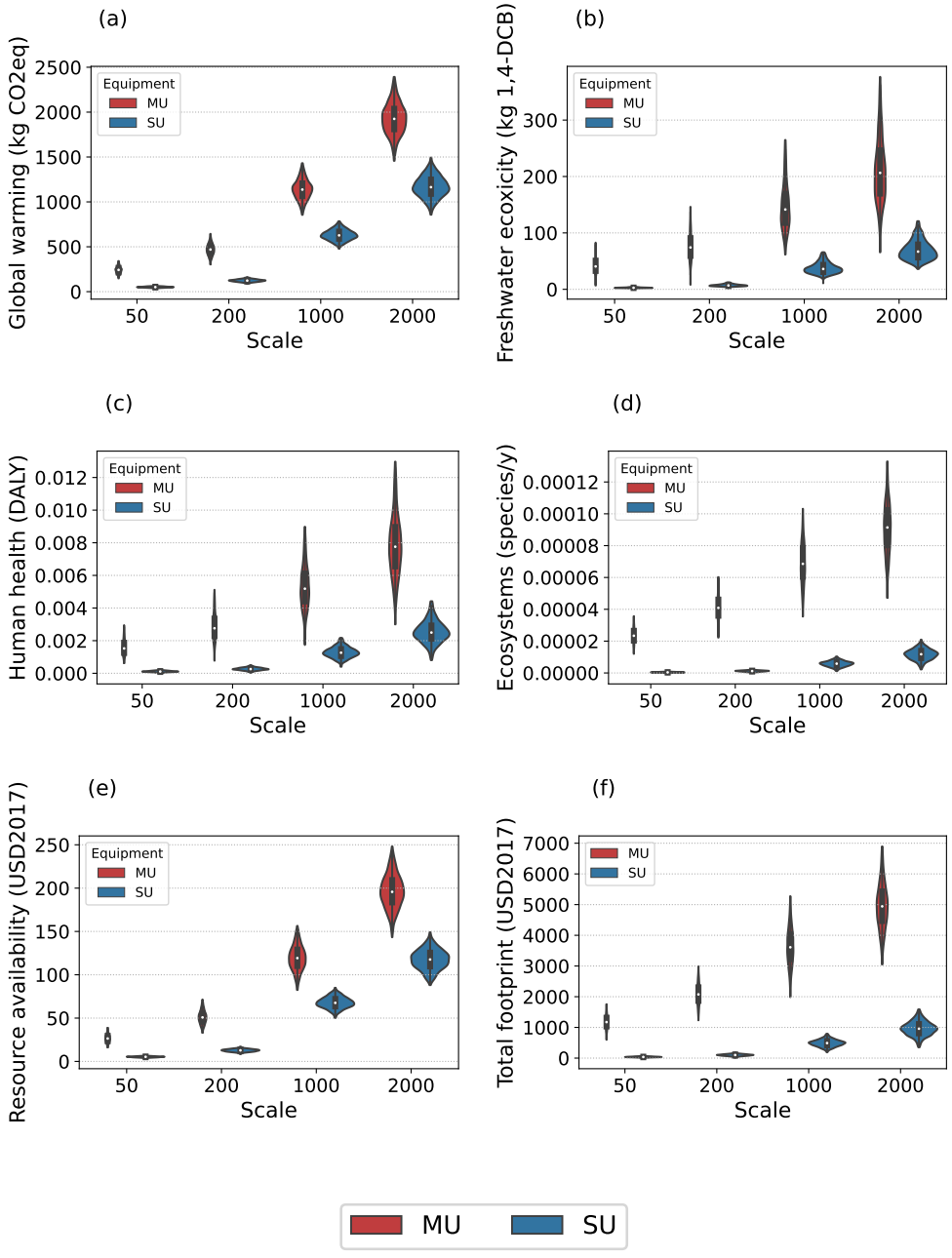


Figure 7.9: Uncertainty quantification for representative environmental indicators for different manufacturing scales a and equipment type e: (a) global warming midpoint, (b) freshwater ecotoxicity midpoint, (c) human health endpoint, (d) ecosystems endpoint, (e) resource availability endpoint and (f) total monetised footprint

As can be seen in Figure 7.9a-f, the uncertainty ranges for MU indicate that MU processes are probabilistically more impactful that SU. In the case of the global warming midpoint (Figure 7.9a) and resource availability endpoint (Figure 7.9e) the variability is relatively similar for both SU and MU process. This is due to the dependence of these indicators on electricity usage, which is lower for SU, whilst assuming the same order of magnitude of variability compared to the MU process. Instead, the uncertainty in the inventory of phosphoric acid in MU manufacturing propagates to most indicators and determines a relatively large uncertainty range (Figure 7.9b, c, d and f). This creates some overlap with the SU footprints, thus indicating a potentially comparable footprint for some indicators, for instance human health (Figure 7.9c).

Overall, despite this large uncertainty introduced by phosphoric acid, MU consistently results in larger footprints compared to SU (Figure 7.9f). Given that each range was obtained through independent sampling, one can be certain of the relative difference only through the implementation of dependent sampling in the OpenLCA platform, thus accounting for potential interactions and correlations between processes that use the same resources (i.e. background inventories). In addition to this, further complexities arise from the integration of process uncertainty generated via independent sampling of the techno-economic models and how to reconcile them with those generated with the LCA platform. Once an integrated cost, throughput and environmental uncertainty space is quantified, data sets of uncertainty realisations could be integrated within a more comprehensive *uncertainty-aware* capacity planning tool.

7.4 Conclusions

This Chapter presents the development of deterministic models for capacity planning integrating environmental sustainability as an additional KPI within the optimisation. The framework relies on an established LCA methodology considering midpoint and endpoint impacts and allowing for a quantification of the underlying uncertainty in optimisation inputs. The results highlighted the key material contributors within the context of SU and MU manufacturing, highlighting that SU processes are more sustainable than MU processes because of reductions in energy usage, water and solvents for intrabatch cleaning. The uncertainty analysis has also reinforced these observations alongside the optimi-

sation results. Overall, it emerged that a key driver in the reduction of environmental footprint is the minimisation of resource use per unit product.

This can be achieved in several ways. Firstly, through the adoption of SU equipment as shown in this Chapter. Secondly, through the integration of uncertainty in supply and demand within the optimisation to prepare to meet target demands and avoid sub-optimal planning, requiring additional capacity to be installed, thus resulting in additional emissions and cost. This would correspond to sub-optimal decisions with respect to the optimisation problem under quantified process and demand uncertainty. Hence, the frameworks presented in Chapters 3-5 are also expected to improve overall sustainability of the supply chains, alongside the said objectives of economic benefits and therapy availability under uncertainty. Ultimately, a unified framework to quantify cost, throughput, environmental footprints uncertainties is fundamental to then perform optimisation under uncertainty and assist decision-making. The output of the integrated framework corresponds to *uncertainty-aware* investment plans which perform better in real-life uncertain scenarios compared to deterministic solutions with respect to *economic*, *environmental* and *social* sustainability.

Chapter 8

Conclusions & outlook

8.1 Summary of contributions

This thesis presents the development of a systematic framework for investment planning of pharmaceutical supply chains. Current industrial trends highlight a growing need for tools that help assess the resilience and sustainability of investments *a priori*. In this space, computer-aided tools provide assessment and optimisation capabilities that help assess alternatives and de-risk rapid scale up decisions in emerging pharmaceutical sectors.

The review of literature with respect to existing planning tools presented in Chapter 2 highlights that currently frameworks do not consider the inherent process uncertainty that emerges in real life scenarios and a pressing need to scale-up requires investment decisions to be made before process capabilities are known for certain and optimised. The problem of planning under demand uncertainty has been extensively studied. However, knowledge gaps emerge with respect to integrating early-stage process uncertainty in cases where capacity needs rapidly repurposing to react to unforeseen demand fluctuations and in cases where investments can be planned in advanced based on future demand predictions. In addition to this, there emerges an opportunity to quantify sustainability metrics of candidate optimisation solutions, thereby informing decision-makers of the footprints of each investment planning decision.

To this end, the developed framework aims to quantify and optimise supply chain performance. First, techno-economic modelling and life cycle analysis are performed to economic, environmental and throughput-related indicators. Second, an optimisation model for supply chain planning is developed, based on input data generated in the quantification step. The outputs of the optimisation serve as decision-support to screen alternatives and assess the impact of input parameters on network performance.

The quantification of uncertainty and its integration in the optimisation is crucial to make the decision-support tool outputs *uncertainty-aware*. Chapter 3 focuses on the quantification of underlying process-specific manufacturing uncertainties via uncertainty analysis and global sensitivity analysis. It is found the process uncertainty is process-specific and the observed variance can be attributed to a set of recurring critical inputs. These include upstream process (USP) productivity, recovery of chromatography steps in downstream process (DSP) and the assumed cost of raw materials, including genetic components. A

strong volumetric increase is recorded for batch size, whereas other output ranges remain equivalent across scales.

In Chapter 4, the generated process parameter distributions are integrated within the optimisation with a comparative assessment of robust and stochastic approaches in two-stage supply chain design problems. Worst-case deterministic optimisation yields investments which result in a 10-fold increase in supply chain costs compared to less conservative decisions obtained through stochastic programming. The stochastic simulation assessment step serves as method to check the simulated performance the identified network design and quantify a probabilistic service level. Stochastic designs are found to vary based on the input samples considered for scenario-representation of uncertainty; however, their performance is consistent and comparative to deterministic worst-case designs with simulated service levels $> 95\%$ if based on a scenario sets of > 16 samples form the original uncertainty dataset. Overall, compared to deterministic designs, the results highlight that process uncertainty propagates to investment planning decisions which justifies its integration in the framework.

As not all uncertainties can be quantified *a priori*, Chapter 5 explores the impact of an unknown unknown on the performance of network designs. As an example, we consider the realisation of a sudden demand for vaccines, which has not been accounted for upon optimisation, and explore design flexibility. Results help delineate a feasibility space of a given design which indicates a clear trade-off between prioritising the production of the starting product and the new product introduction. Effectively, the optimisation formulation accounts for the underlying process uncertainty with respect to each product. We prove that the scenario-based representation of the uncertainty space remains accurate for a fixed design. This allows the quantification of an *uncertainty-aware* flexibility of a design in meeting two competing demands and serves as metric to assess the real capabilities of an asset in catering for multiple products under development.

To address problems where all demands can be anticipated with probabilistic estimates, Chapter 6 presents a multi-stage stochastic problem, which considers that the previously quantified process uncertainty realises along the time horizon alongside demand uncertainty. By comparing different facility design and equipment decisions, we highlight the better performance of re-configurable SU manufacturing configurations in rapidly adapting to known uncertainty, which is expected to hold true for unknown unknowns too.

Finally, Chapter 7 focuses on the quantification of sustainability metrics through an LCA methodology and its integration in the supply chain optimisation framework. Trade-offs between cost, environmental and social indicators are explored. It emerges that reconfigurable Single-use (SU)-based facilities result in smaller environmental footprints compared to fixed stainless steel multi-use (MU) facilities. This is primarily due the extensive use of phosphoric acid in cleaning steps and a larger use of electricity. SU manufacturing also results to be the most cost-effective option. The key trade-off relates to the social benefit of supplying medicines and the impact on the environment.

Next, the main conclusions that result from the findings of this thesis are outlined (Section 8.2), followed by a discussion on emerging research avenues that motivate further work in this area (Section 8.3).

8.2 Main conclusions

The main conclusions that can be drawn from the studies presented in Chapter 3, 4, 5, 7 are listed below:

1. Investment planning for emerging biopharmaceutical products is best performed with considerations on process uncertainty. Through the quantification of process-specific and scale-dependent process uncertainties we have demonstrated that the *uncertainty-aware* network designs perform better under simulated uncertainty compared to deterministic designs that disregard uncertainty in process capabilities for platforms under development. This is crucial for next-generation supply chains which are continuously undergoing scale-up, integrating new product approvals and catering for growing demands before future intensified process capabilities are known.

2. Network costs increase with increasing target levels of reliability. Facilities are progressively overdesigned as the worst-case realisations of uncertainties become increasingly accounted for. This is not unexpected however there emerges a trade-off between the achievable supply reliability with respect to different products in networks with resource constraints. This requires decision-makers to make decisions on prioritising either products.

3. Not all uncertainties related to manufacturing, supply chain operations and demands can be quantified *a priori*. Hence, there is a limit to the level of

uncertainty-awareness that one can embed in optimisation design decisions. It is expected that the establishment of designs which allow for short-term recourse actions are likely going to result in both better performance in responding to unforeseen uncertainties (unknown unknowns) in real-life scenarios. This highlights the potential of modular SU manufacturing in improving supply chain resilience to uncertainty and overall make supply chain capacity planning decisions more responsive to a range of long-term and short-term uncertainties.

4. Modular SU manufacturing promises to be the most cost-effective option in industry and our findings agree with this trend. In addition to this, the LCA studies highlight that SU manufacturing results in lower impacts than MU manufacturing. Thus, in this case the *economic* and *environmental* sustainability are not conflicting objectives. Given that cost-effectiveness in real life scenarios is linked to implementing solutions that mitigate adverse impacts of uncertainty, more comprehensive and reliable trade-offs assessment require an integration of sustainability and the uncertainty in its quantification in the developed optimisation-based frameworks for planning under uncertainty.

8.3 Future directions

This thesis has contributed towards setting the foundation of a comprehensive decision-support tool for pharmaceutical supply chain planning integrating sustainability and a range of uncertainties. At current state, decision-makers in the industry (e.g. production, supply chain, project managers) could use the tools to systematically compare alternatives in terms of process capabilities (Chapter 3), establish manufacturing capacity that mitigates the risk of shortages (Chapter 4), explore its flexibility (Chapter 5) and deterministically compare designs with respect to cost and environmental footprints (Chapter 7).

The suggestions discussed in following sections focus on (i) integrating process uncertainty and demand uncertainty in multi-period contexts, further developing the framework presented in (Chapter 6), (ii) integrating sustainability-related uncertainty for more reliable comparative assessments, aiming to create unified frameworks.

8.3.1 Investment planning

In Chapter 6, we demonstrate the development of a multi-stage stochastic planning problem which considers demand uncertainty and integrates a consideration of process uncertainty. Process uncertainty is integrated with a deterministic appproach, keeping the model tractable. Despite this deterministic description, the resulting problem size requires an optimality gap threshold of 10% to ensure that results are obtained within reasonable timeframes.

It is worth investigating the potential of planning for shorter look-ahead time-periods to reduce multi-stage scenario tree size. In doing so, one could consider the solution of the entire time-horizon with a rolling horizon approach and compare the solution. It would be interesting to check whether solutions change significantly both for the MU and SU case. A reduction in the time horizon could allow an integration with a two-stage scenario-based approach of process uncertainty in the first stages of the rolling horizon, whilst still assuming that process capabilities become known as time progresses. It would be valuable to compare the resulting solutions with the ones obtained with a multi-stage robust approach and study the impact of assumed process realisations. Robust optimisation is expected to achieve solutions that balance the conservativeness of worst-case deterministic optimisation and stochastic programming.

8.3.2 Flexibility metrics quantification

Chapter 5 has explored the concept of flexibility quantification for a given supply chain design. Using the quantified process uncertainty space, it proposes the concept of an *uncertainty-aware* flexibility which quantifies a product supply reliability metric for each solution point that forms the space.

First, it is worth trying to solve the space with increased granularity. Given that n scenarios are used to approximate the probability space and the chance constraints, there is a limit to the achievable granularity of the grid to intervals of $\frac{1}{n}$. This creates scope for exploring alternative chance constraint formulations which simplify the underlying process uncertainty to key parameters (e.g. batch size) and reformulate chance constraints deterministically with a percentile approach. The accuracy of this reformulation could be compared with the scenario-based chance constraints and indicate the path forward. Potentially, simplifying the chance constraint could help improve the scalability of the problem to more

products and larger network cases.

Second, the *uncertainty-aware* flexibility has yet to be compared to the solution obtained with a feasibility problem that is solved assuming deterministic process capabilities. It would be expected that the feasible region would be optimistically larger if one does not consider uncertainty. Effectively, this corresponds to the flexibility that can be quantified when process uncertainty is reduced, for instance in future stages of supply chain operation.

Third, feasibility assessments of network designs are very valuable to quantify flexibility metrics to *unknown* unknowns. It would be interesting to consider a range of unforeseen uncertainties including disruptions, plant shutdowns and multiple unforeseen demands. Ultimately, once *good* flexible designs are identified, their performance could be simulated with discrete-event simulation approaches, perhaps generating outcomes of uncertainties the design is *aware* of, as well as unforeseen uncertainties.

8.3.3 *Uncertainty-aware* and sustainable designs

In Chapter 3 cost and throughput-uncertainty is quantified. This is integrated in the optimisation in Chapters 4, 5 and 6 through different strategies which are compared and assessed. In Chapter 7, sustainability metrics are quantified both as nominal point outputs and as uncertainty ranges, whilst we only illustrate a deterministic framework. Ultimately, supply chain performance is best optimised if uncertainty is considered within the optimisation.

Before integration, one would need to reconcile the sustainability uncertainty, with cost and throughput-uncertainty. All these analyses were performed independently; however, the actual independence of all inputs and amongst processes considered in the techno-economic and LCA models would need to be carefully assessed and the uncertainty analysis sampling strategy structured accordingly. Once uncertainties are obtained, a comprehensive optimisation model would need to embed constraints which compute cost, environmental and social indicators and integrate input uncertainty whether through stochastic or robust optimisation. The tool would allow the identification of *uncertainty-aware* designs which minimise cost, environmental footprint and maximising therapy availability.

8.3.4 Other applications

We focused on viral vector applications to illustrate how decision-support can tackle some emerging supply chain industrial challenges. The frameworks presented in this work are transferable to any product application of interest. With supply chain resilience and sustainability becoming growing priorities across the entire pharmaceutical sector, it would be interesting to explore the application of the developed tools to other advanced therapeutic product applications and supply chain case studies, so as to identify case-specific trends and trade-offs.

Dissemination of outputs

Journal Papers

10. **Sarkis, M.**, Fyfe, A., Shah, N. & Papathanasiou, M.M. Towards 3-fold sustainability in biopharmaceutical supply chain planning: single-use vs stainless steel manufacturing. *To be submitted to ACS Sustainable Engineering*. In preparation.

9. **Sarkis, M.**, Sachio, S., Kontoravdi, C., & Papathanasiou, M.M. Flexibility and sustainability in biopharma: computer-aided platforms for process development and network design, *To be submitted to Computer & Chemical Engineering Special Issue*. In preparation (invited).

8. **Sarkis, M.**, Shah, N. & Papathanasiou, M.M. Resilient multi-product next-generation pharmaceutical supply chains: quantification of flexibility under uncertainty. *Submitted to Journal of Advanced Processing & Manufacturing*. Under review.

7. **Sarkis, M.**, Shah, N. & Papathanasiou, M.M. Resilient pharmaceutical supply chain planning: assessment of strategies for process uncertainty integration in network design problems. *Submitted to Computers & Chemical Engineering*. Under review.

6. **Sarkis, M.**, Fyfe, A., Kontoravdi, C. & Papathanasiou, M.M. (2024). Towards a Net Zero, socially sustainable and eco-efficient biopharma: how far are we?, *Current Opinion in Chemical Engineering* 44, 101027.

5. Triantafyllou, N., **Sarkis, M.**, Krassakopoulou, A., Shah, N. & Papathanasiou, M.M., & Kontoravdi, C. (2024). Uncertainty quantification for gene delivery methods: a roadmap for pDNA manufacturing from phase I clinical trials to commercialisation, *Biotechnology Journal*, 19:2300103.

4. **Sarkis, M.**, Shah, N. & Papathanasiou, M.M. (2023). Characterization of key manufacturing uncertainties in next-generation therapeutics and vaccines across scales, *Journal of Advanced Manufacturing and Processing* 5(3), e10158.

3. Bernardi, A., **Sarkis, M.**, Triantafyllou, N., Shah, N. & Papathanasiou, M.M. (2022). Assessment of intermediate storage and distribution in personalised medicine, *Computers & Chemical Engineering* 157, 107582.

2. **Sarkis, M.**, Bernardi, A., Shah, N. & Papathanasiou, M.M. (2021). Decision support tools for next-generation vaccines and advanced therapy medicinal products: present and future, *Current Opinion in Chemical Engineering* 32, 100689.

1. **Sarkis, M.**, Bernardi, A., Shah, N. & Papathanasiou, M.M. (2021). Emerging challenges and opportunities in pharmaceutical manufacturing and distribution', *Processes* 9(3), 457.

Peer-reviewed Conference Papers

4. **Sarkis, M.**, Shah, N. & Papathanasiou, M.M.. Integrating process and demand uncertainty in capacity planning for next-generation pharmaceutical supply chains, *To be submitted to Book of Abstracts of the 35th European Symposium on Computer Aided Process Engineering*, In preparation.

3. **Sarkis, M.**, Sachio, S., Kontoravdi, C., Shah, N. & Papathanasiou, M.M. (2022). Towards 3-fold sustainability in biopharmaceutical process development and product distribution, *Systems and Control Transactions*, 3, 75-83.

2. **Sarkis, M.**, Bernardi, A., Shah, N. & Papathanasiou, M.M. (2022). Integrating environmental sustainability in next-generation biopharmaceutical supply chains, *Book of Abstracts of the 33rd European Symposium on Computer Aided Process Engineering*, 1633-1638.

1. **Sarkis, M.**, Tak, K., Chachuat, B., Shah, N. & Papathanasiou, M.M. (2021). Towards resilience in next-generation vaccines and therapeutics

supply chains, *Book of Abstracts of the 32nd European Symposium on Computer Aided Process Engineering*, 931-936.

Book Chapters

1. Triantafyllou, N., **Sarkis, M.**, Shah, N., Papathanasiou, M.M. & Kontoravdi, C. (2024). *Integrated process and supply chain design and optimisation*, Cell Engineering 11.

Bibliography

A. M. Aguirre, S. Liu, and L. G. Papageorgiou. Optimisation approaches for supply chain planning and scheduling under demand uncertainty. *Chemical Engineering Research and Design*, 138:341–357, 10 2018. ISSN 02638762. doi: 10.1016/j.cherd.2018.08.021.

A. Almansoori and N. Shah. Design and operation of a future hydrogen supply chain: Snapshot model. *Chemical Engineering Research and Design*, 84(6 A): 423–438, 2006. ISSN 02638762. doi: 10.1205/cherd.05193.

A. Almansoori and N. Shah. Design and operation of a future hydrogen supply chain: Multi-period model. *International Journal of Hydrogen Energy*, 34(19): 7883–7897, 10 2009. ISSN 03603199. doi: 10.1016/j.ijhydene.2009.07.109.

A. C. Amaro and A. P. Barbosa-Póvoa. Planning and scheduling of industrial supply chains with reverse flows: A real pharmaceutical case study. *Computers and Chemical Engineering*, 32(11):2606–2625, 11 2008. ISSN 00981354. doi: 10.1016/j.compchemeng.2008.03.006.

A. C. Amaro and A. P. Barbosa-Póvoa. The effect of uncertainty on the optimal closed-loop supply chain planning under different partnerships structure. *Computers and Chemical Engineering*, 33(12):2144–2158, 12 2009. ISSN 00981354. doi: 10.1016/j.compchemeng.2009.06.003.

S. Ansorge, S. Lanthier, J. Transfiguracion, Y. Durocher, O. Henry, and A. Kamen. Development of a scalable process for high-yield lentiviral vector production by transient transfection of HEK293 suspension cultures. *Journal of Gene Medicine*, 11(10):868–876, 2009. ISSN 1099498X. doi: 10.1002/jgm.1370.

F. Arabi, V. Mansouri, and N. Ahmadbeigi. Gene therapy clinical trials, where do we go? An overview, 9 2022. ISSN 19506007.

AstraZeneca. Making the COVID-19 Vaccine: Fact Sheet. 2021.

AstraZeneca. Seven pharma CEOs announce new joint action to accelerate net zero healthcare, 2022. URL https://www.astrazeneca.com/media-centre/articles/2022/seven-pharma-ceos-announce-new-joint-action-to-accelerate-net-zero-healthcare.html.

O. Badejo and M. Ierapetritou. Mathematical Programming Approach to Optimize Tactical and Operational Supply Chain Decisions under Disruptions. *Industrial and Engineering Chemistry Research*, 61(45):16747–16763, 11 2022. ISSN 15205045. doi: 10.1021/acs.iecr.2c01641.

O. Badejo and M. Ierapetritou. Enhancing pharmaceutical supply chain resilience: A multi-objective study with disruption management. *Computers and Chemical Engineering*, 188, 9 2024. ISSN 00981354. doi: 10.1016/j.compchemeng.2024.108769.

J. Balasubramanian and I. E. Grossmann. Approximation to Multistage Stochastic Optimization in Multiperiod Batch Plant Scheduling under Demand Uncertainty. 2004. doi: 10.1021/ie030308. URL https://pubs.acs.org/sharingguidelines.

A. P. Barbosa-Póvoa. Process supply chains management - where are we? Where to go next? *Frontiers in Energy Research*, 2(JUN), 2014. ISSN 2296598X. doi: 10.3389/fenrg.2014.00023.

A. P. Barbosa-Póvoa, C. da Silva, and A. Carvalho. Opportunities and challenges in sustainable supply chain: An operations research perspective, 7 2018. ISSN 03772217.

M. Bauler, J. K. Roberts, C. C. Wu, B. Fan, F. Ferrara, B. H. Yip, S. Diao, Y. I. Kim, J. Moore, S. Zhou, M. M. Wielgosz, B. Ryu, and R. E. Throm. Production of Lentiviral Vectors Using Suspension Cells Grown in Serum-free Media. *Molecular Therapy - Methods and Clinical Development*, 17:58–68, 6 2020a. ISSN 23290501. doi: 10.1016/j.omtm.2019.11.011.

M. Bauler, J. K. Roberts, C. C. Wu, B. Fan, F. Ferrara, B. H. Yip, S. Diao, Y. I. Kim, J. Moore, S. Zhou, M. M. Wielgosz, B. Ryu, and R. E. Throm. Production of Lentiviral Vectors Using Suspension Cells Grown in Serum-free

Media. *Molecular Therapy - Methods and Clinical Development*, 17:58–68, 6 2020b. ISSN 23290501. doi: 10.1016/j.omtm.2019.11.011.

A. Bernardi, M. Papathanasiou, M. Lakelin, and N. Shah. Assessment of intermediate storage and distribution nodes in personalised medicine. 2021.

A. Bonfill, A. Espuña, and L. Puigjaner. Addressing robustness in scheduling batch processes with uncertain operation times. *Industrial and Engineering Chemistry Research*, 44(5):1524–1534, 3 2005. ISSN 08885885. doi: 10.1021/ie049732g.

S. Bose and J. F. Pekny. Engineering A model predictive framework for planning and scheduling problems: a case study of consumer goods supply chain. *Computers & Chemical Computers and Chemical Engineering*, 24, 2000. URL `www.elsevier.com/locate/compchemeng`.

K. Budzinski, D. Constable, D. D'Aquila, P. Smith, S. R. Madabhushi, A. Whiting, T. Costelloe, and M. Collins. Streamlined life cycle assessment of single use technologies in biopharmaceutical manufacture. *New Biotechnology*, 68: 28–36, 5 2022. ISSN 18764347. doi: 10.1016/j.nbt.2022.01.002.

J. T. Bulcha, Y. Wang, H. Ma, P. W. L. Tai, and G. Gao. Viral vector platforms within the gene therapy landscape. *Signal Transduction and Targeted Therapy*, 6(53):1–23, 12 2021. ISSN 2059-3635. doi: 10.1038/s41392-021-00487-6.

P. Bunnak, R. Allmendinger, S. V. Ramasamy, P. Lettieri, and N. J. Titchener-Hooker. Life-cycle and cost of goods assessment of fed-batch and perfusion-based manufacturing processes for mAbs. *Biotechnology Progress*, 32(5): 1324–1335, 9 2016. ISSN 15206033. doi: 10.1002/btpr.2323.

K. J. Caldwell, S. Gottschalk, and A. C. Talleur. Allogeneic CAR Cell Therapy—More Than a Pipe Dream, 1 2021. ISSN 16643224.

E. Cameau, A. Pedregal, and C. Glover. Cost modelling comparison of adherent multi-trays with suspension and fixed-bed bioreactors for the manufacturing of gene therapy products. *Cell and Gene Therapy Insights*, 5(11):1663–1674, 11 2019. ISSN 20597800. doi: 10.18609/cgti.2019.175.

E. Capra, A. Gennari, A. Loche, and C. Temps. Viral-vector therapies at scale: today's challenges and future opportunities, 2022.

S. R. Cardoso, A. P. F. Barbosa-Póvoa, and S. Relvas. Design and planning of supply chains with integration of reverse logistics activities under demand uncertainty. *European Journal of Operational Research*, 226(3):436–451, 5 2013. ISSN 03772217. doi: 10.1016/j.ejor.2012.11.035.

P. Castro, A. P. Barbosa-Póvoa, and H. Matos. An improved RTN continuous-time formulation for the short-term scheduling of multipurpose batch plants. *Industrial and Engineering Chemistry Research*, 40(9):2059–2068, 5 2001. ISSN 08885885. doi: 10.1021/ie000683r.

P. M. Castro, I. Harjunkoski, and I. E. Grossmann. Optimal short-term scheduling of large-scale multistage batch plants. *Industrial and Engineering Chemistry Research*, 48(24):11002–11016, 2009. ISSN 08885885. doi: 10.1021/ie900734x.

Catapult. Capability and Capacity Analysis Cell and Gene Therapy GMP Manufacturing in the UK. Technical report, 2023.

S. Cecchini, T. Virag, and R. M. Kotin. Reproducible high yields of recombinant adeno-associated virus produced using invertebrate cells in 0.02-to 200-liter cultures. *Human Gene Therapy*, 22(8):1021–1030, 8 2011. ISSN 10430342. doi: 10.1089/hum.2010.250.

A. Charnes and W. W. Cooper. Chance-Constrained Programming. Technical Report 1, 1959. URL `https://about.jstor.org/terms`.

Y. Chen, L. Mockus, S. Orcun, and G. V. Reklaitis. Simulation-optimization approach to clinical trial supply chain management with demand scenario forecast. *Computers and Chemical Engineering*, 40:82–96, 5 2012. ISSN 00981354. doi: 10.1016/j.compchemeng.2012.01.007.

CHMP. Package Insert - LUXTURNA (voretigene neparvovec-rzyl). Technical report, 2018. URL `www.fda.gov/medwatch`.

CHMP. Package Insert: KYMRIAH (tisagenlecleucel). 2019a.

CHMP. Package Insert: ZONGELSMA (onasemnogene abeparvovec-xioi). 2019b. URL `www.fda.gov/medwatch`.

CHMP. Package Insert: COVID-19 Vaccine AstraZeneca, COVID 19 Vaccine (ChAdOx1 S [recombinant]). 2021a.

CHMP. Package Insert: Jansenn COVID-19 Vaccine. 2021b. URL `https://www.cdc.gov/vaccines/covid-19/clinical-considerations/managing-anaphylaxis.html`.

M. Colvin and C. T. Maravelias. A stochastic programming approach for clinical trial planning in new drug development. *Computers and Chemical Engineering*, 32(11):2626–2642, 11 2008. ISSN 00981354. doi: 10.1016/j.compchemeng.2007.11.010.

M. Colvin and C. T. Maravelias. Scheduling of testing tasks and resource planning in new product development using stochastic programming. *Computers and Chemical Engineering*, 33(5):964–976, 5 2009. ISSN 00981354. doi: 10.1016/j.compchemeng.2008.09.010.

R. M. Comisel, B. Kara, F. H. Fiesser, and S. S. Farid. Gene therapy process change evaluation framework: Transient transfection and stable producer cell line comparison. *Biochemical Engineering Journal*, 176, 12 2021a. ISSN 1873295X. doi: 10.1016/j.bej.2021.108202.

R. M. Comisel, B. Kara, F. H. Fiesser, and S. S. Farid. Lentiviral vector bioprocess economics for cell and gene therapy commercialization. *Biochemical Engineering Journal*, 167(107868), 3 2021b. ISSN 1873295X. doi: 10.1016/j.bej.2020.107868.

Cytiva. Xcellerex: cell culture bioreactor systems. Technical report.

Cytiva. Strategies to address the viral vector manufacturing shortage, 2021. URL `https://www.cytivalifesciences.com/en/us/solutions/bioprocessing/knowledge-center/strategies-for-viral-vector-manufacturing`.

C. Ding and M. Ierapetritou. A novel framework of surrogate-based feasibility analysis for establishing design space of twin-column continuous chromatography. *International Journal of Pharmaceutics*, 609, 11 2021. ISSN 18733476. doi: 10.1016/j.ijpharm.2021.121161.

I. Duarte, B. Mota, T. Pinto-Varela, and A. P. Barbosa-Póvoa. Pharmaceutical industry supply chains: How to sustainably improve access to vaccines? *Chemical Engineering Research and Design*, 182:324–341, 6 2022. ISSN 02638762. doi: 10.1016/j.cherd.2022.04.001.

B. A. Elias Mota, A. I. Cerqueira de Sousa Gouveia Carvalho, M. I. Azevedo Rodrigues Gomes, and A. P. Ferreira Dias Barbosa-Povoa. Business strategy for sustainable development: Impact of life cycle inventory and life cycle impact assessment steps in supply chain design and planning. *Business Strategy and the Environment*, 29(1):87–117, 1 2020. ISSN 10990836. doi: 10.1002/bse.2352.

EMA. Glybera, INN - alipogene tiparvovec. Technical report, 2015. URL `www.ema.europa.eu/contact`.

EMA. Strimvelis - autologous CD34+ enriched cell fraction that contains CD34+ cells transduced with retroviral vector that encodes for the human ADA cDNA sequence. Technical report, 2016. URL `www.ema.europa.eu/contact`.

EMA. Advanced therapy medicinal products: Overview, 2020. URL `https://www.ema.europa.eu/en/human-regulatory/overview/advanced-therapy-medicinal-products-overview`.

EMA. AstraZeneca's COVID-19 vaccine: EMA finds possible link to very rare cases of unusual blood clots with low blood platelets, 2021. URL `https://www.ema.europa.eu/en/news/astrazenecas-covid-19-vaccine-ema-finds-possible-link-very-rare-cases-unusual-blood-clots-low-blood-platelets`.

O. Farnós, E. Gelaye, K. Trabelsi, A. Bernier, K. Subramani, H. Kallel, M. Yami, and A. A. Kamen. Establishing a robust manufacturing platform for recombinant veterinary vaccines: An adenovirus-vector vaccine to control newcastle disease virus infections of poultry in Sub-Saharan Africa. *Vaccines*, 8(2):1–24, 6 2020. ISSN 2076393X. doi: 10.3390/vaccines8020338.

S. Fedosyuk, T. Merritt, M. P. Peralta-Alvarez, S. J. Morris, A. Lam, N. Laroudie, A. Kangokar, D. Wright, G. M. Warimwe, P. Angell-Manning, A. J. Ritchie, S. C. Gilbert, A. Xenopoulos, A. Boumlic, and A. D. Douglas. Simian adenovirus vector production for early-phase clinical trials: A simple

method applicable to multiple serotypes and using entirely disposable product-contact components. *Vaccine*, 37(47):6951–6961, 11 2019. ISSN 18732518. doi: 10.1016/j.vaccine.2019.04.056.

R. G. Ferreira, N. F. Gordon, R. Stock, and D. Petrides. Adenoviral vector covid-19 vaccines: Process and cost analysis. *Processes*, 9(8), 8 2021. ISSN 22279717. doi: 10.3390/pr9081430.

Gardner B. Oxford Covid vaccine 'bottlenecks' mean UK's full order will not be ready until end of next year, 11 2020.

G. Gatica, L. G. Papageorgiou, and N. Shah. Capacity planning under uncertainty for the pharmaceutical industry. *Chemical Engineering Research and Design*, 81(A):665–678, 2003a. ISSN 02638762. doi: 10.1205/026387603322150516.

G. Gatica, L. G. Papageorgiou, and N. Shah. An Aggregation Approach For Capacity Planning Under Uncertainty for the Pharmaceutical Industry. 2003b.

G. P. Georgiadis and M. C. Georgiadis. Optimal planning of the COVID-19 vaccine supply chain. *Vaccine*, 39(37):5302–5312, 8 2021. ISSN 18732518. doi: 10.1016/j.vaccine.2021.07.068.

S. Ghosh, A. M. Brown, C. Jenkins, and K. Campbell. Viral Vector Systems for Gene Therapy: A Comprehensive Literature Review of Progress and Biosafety Challenges, 3 2020. ISSN 24701246.

J. Gjerdrum, N. Shah, and L. G. Papageorgiou. A combined optimization and agent-based approach to supply chain modelling and performance assessment. *Production Planning and Control*, 12(1):81–88, 1 2001. ISSN 09537287. doi: 10.1080/09537280150204013.

Global Market Insights. Gene Therapy Market, 2023.

E. Gökalp and J. Branke. Pharmaceutical R & D pipeline management under trial duration uncertainty. *Computers and Chemical Engineering*, 136, 5 2020. ISSN 00981354. doi: 10.1016/j.compchemeng.2020.106782.

Grand View Research. Gene Therapy Market Size & Trends, 2023.

J. C. Grieger, S. M. Soltys, and R. J. Samulski. Production of recombinant adeno-associated virus vectors using suspension HEK293 cells and continuous harvest of vector from the culture media for GMP FIX and FLT1 clinical vector. *Molecular Therapy*, 24(2):287–297, 2 2016. ISSN 15250024. doi: 10.1038/mt.2015.187.

I. Grossmann. Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE Journal*, 51(7):1846–1857, 7 2005. ISSN 00011541. doi: 10.1002/aic.10617.

I. Grossmann. Advances in Mathematical Programming Models for Enterprise-Wide Optimization. 2012. URL `http://www.cheme.cmu.edu/focapo`.

G. Guillén-Gosálbez, F. D. Mele, and I. E. Grossmann. A bi-criterion optimization approach for the design and planning of hydrogen supply chains for vehicle use. *AIChE Journal*, 56(3):650–667, 3 2010. ISSN 00011541. doi: 10.1002/aic.12024.

A. Gupta and C. D. Maranas. A two-stage modeling and solution framework for multisite midterm planning under demand uncertainty. *Industrial and Engineering Chemistry Research*, 39(10):3799–3813, 2000. ISSN 08885885. doi: 10.1021/ie9909284.

A. Gupta, C. D. Maranas, C. M. Mcdonald, and D. Lycra. Midterm Supply Chain Planning Under Demand Uncertainty : Customer Demand Satisfaction and Inventory Management. 2003.

P. Gupta, N. Kateja, S. Mishra, H. Kaur, and A. S. Rathore. Economic assessment of continuous processing for manufacturing of biotherapeutics. *Biotechnology Progress*, 37(2), 3 2021. ISSN 15206033. doi: 10.1002/btpr.3108.

HM Government. Life Sciences Vision. Technical report, 2021.

A. Hugo and E. N. Pistikopoulos. Environmentally conscious long-range planning and design of supply chain networks. In *Journal of Cleaner Production*, volume 13, pages 1471–1491. Elsevier Ltd, 2005. doi: 10.1016/j.jclepro.2005.04.011.

W. Y. Hung, N. J. Samsatli, and N. Shah. Object-oriented dynamic supply-chain modelling incorporated with production scheduling. In *European Journal of*

Operational Research, volume 169, pages 1064–1076, 3 2006. doi: 10.1016/j.ejor.2005.02.009.

D. Ibrahim, Z. Kis, K. Tak, M. M. Papathanasiou, C. Kontoravdi, B. Chachuat, and N. Shah. Model-based planning and delivery of mass vaccination campaigns against infectious disease: Application to the covid-19 pandemic in the UK. *Vaccines*, 9(12), 12 2021. ISSN 2076393X. doi: 10.3390/vaccines9121460.

M. G. Ierapetritou and C. A. Floudas. Effective continuous-time formulation for short-term scheduling. 1. Multipurpose batch processes. *Industrial and Engineering Chemistry Research*, 37(11):4341–4359, 1998. ISSN 08885885. doi: 10.1021/ie970927g.

Intelligen. SuperPro Designer (v12).

G. Jagschies, E. Lindskog, K. Lacki, and P. Galliher. *Biopharmaceutical Processing: Development, Design and Implementation of Manufacturing Processes*. Elsevier, 1 edition, 2018.

S. L. Janak, C. A. Floudas, J. Kallrath, and N. Vormbrock. Production scheduling of a large-scale industrial batch plant. II. Reactive scheduling. *Industrial and Engineering Chemistry Research*, 45(25):8253–8269, 12 2006. ISSN 08885885. doi: 10.1021/ie0600590.

S. L. Janak, X. Lin, and C. A. Floudas. A new robust optimization approach for scheduling under uncertainty. II. Uncertainty with known probability distribution. *Computers and Chemical Engineering*, 31(3):171–195, 1 2007. ISSN 00981354. doi: 10.1016/j.compchemeng.2006.05.035.

M. J. Jenkins and S. S. Farid. Cost-effective bioprocess design for the manufacture of allogeneic CAR-T cell therapies using a decisional tool with multi-attribute decision-making analysis. *Biochemical Engineering Journal*, 137: 192–204, 9 2018. ISSN 1873295X. doi: 10.1016/j.bej.2018.05.014.

S. Kabra, M. A. Shaik, and A. S. Rathore. Multi-period scheduling of a multi-Stage multi-Product bio-pharmaceutical process. *Computers and Chemical Engineering*, 57:95–103, 10 2013. ISSN 00981354. doi: 10.1016/j.compchemeng.2013.03.009.

M. Kaut. Scenario generation by selection from historical data. *Computational Management Science*, 18(3):411–429, 7 2021. ISSN 16196988. doi: 10.1007/s10287-021-00399-4.

Z. Kis, M. Papathanasiou, R. Calvo-Serrano, C. Kontoravdi, and N. Shah. A model-based quantification of the impact of new manufacturing technologies on developing country vaccine supply chain performance: A Kenyan case study. *Journal of Advanced Manufacturing and Processing*, 1(3), 7 2019a. ISSN 2637-403X. doi: 10.1002/amp2.10025.

Z. Kis, R. Shattock, N. Shah, and C. Kontoravdi. Emerging Technologies for Low-Cost, Rapid Vaccine Manufacture, 1 2019b. ISSN 18607314.

Z. Kis, K. Tak, D. Ibrahim, M. M. Papathanasiou, B. Chachuat, N. Shah, and C. Kontoravdi. Pandemic-response adenoviral vector and RNA vaccine manufacturing. *npj Vaccines*, 7(1), 12 2022. ISSN 20590105. doi: 10.1038/s41541-022-00447-3.

E. Kondili, C. C. Pantelides, and R. W. H. Sargent. A general algorithm for short-term scheduling of batch operations -I. MILP formulation. *Computers and Chemical Engineering*, 17(2):211–227, 1993.

S. Küçükyavuz and R. Jiang. Chance-constrained optimization under limited distributional information: A review of reformulations based on sampling and distributional robustness, 1 2022. ISSN 21924414.

J. M. Laínez, G. Kopanos, A. Espuña, and L. Puigjaner. Flexible design-planning of supply chain networks. *AIChE Journal*, 55(7):1736–1753, 7 2009. ISSN 00011541. doi: 10.1002/aic.11942.

K. Lakhdar, Y. Zhou, J. Savery, N. J. Titchener-Hooker, and L. G. Papageorgiou. Medium term planning of biopharmaceutical manufacture using mathematical programming. *Biotechnology Progress*, 21(5):1478–1489, 9 2005. ISSN 87567938. doi: 10.1021/bp0501571.

K. Lakhdar, S. S. Farid, N. J. Titchener-Hooker, and L. G. Papageorgiou. Medium term planning of biopharmaceutical manufacture with uncertain fermentation titers. *Biotechnology Progress*, 22(6):1630–1636, 11 2006. ISSN 87567938. doi: 10.1021/bp0601950.

K. Lakhdar, J. Savery, L. G. Papageorgiou, and S. S. Farid. Multiobjective long-term planning of biopharmaceutical manufacturing facilities. *Biotechnology Progress*, 23(6):1383–1393, 11 2007. ISSN 87567938. doi: 10.1021/bp0701362.

N. H. Lappas and C. E. Gounaris. Robust optimization for decision-making under endogenous uncertainty. *Computers and Chemical Engineering*, 111: 252–266, 3 2018. ISSN 00981354. doi: 10.1016/j.compchemeng.2018.01.006.

D. Letsios, M. Mistry, and R. Misener. Exact lexicographic scheduling and approximate rescheduling. *European Journal of Operational Research*, 290(2): 469–478, 4 2021. ISSN 03772217. doi: 10.1016/j.ejor.2020.08.032.

B. L. Levine, J. Miskin, K. Wonnacott, and C. Keir. Global Manufacturing of CAR T Cell Therapy, 3 2017. ISSN 23290501.

A. A. Levis and L. G. Papageorgiou. A hierarchical solution approach for multi-site capacity planning under uncertainty in the pharmaceutical industry. *Computers and Chemical Engineering*, 28(5):707–725, 5 2004. ISSN 00981354. doi: 10.1016/j.compchemeng.2004.02.012.

Z. Li and M. G. Ierapetritou. Reactive scheduling using parametric programming. *AIChE Journal*, 54(10):2610–2623, 10 2008. ISSN 00011541. doi: 10.1002/aic.11593.

Z. Li and M. G. Ierapetritou. Rolling horizon based planning and scheduling integration with production capacity consideration. *Chemical Engineering Science*, 65(22):5887–5900, 11 2010. ISSN 00092509. doi: 10.1016/j.ces.2010.08.010.

X. Lin, S. L. Janak, and C. A. Floudas. A new robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty. *Computers and Chemical Engineering*, 28(6-7):1069–1085, 6 2004. ISSN 00981354. doi: 10.1016/j.compchemeng.2003.09.020.

S. Liu and L. G. Papageorgiou. Multiobjective optimisation of production, distribution and capacity planning of global supply chains in the process industry. *Omega (United Kingdom)*, 41(2):369–382, 4 2013. ISSN 03050483. doi: 10.1016/j.omega.2012.03.007.

S. Liu, N. Shah, and L. G. Papageorgiou. Multiechelon supply chain planning with sequence-dependent changeovers and price elasticity of demand under

uncertainty. *AIChE Journal*, 58(11):3390–3403, 11 2012. ISSN 00011541. doi: 10.1002/aic.13732.

S. Liu, A. Yahia, and L. G. Papageorgiou. Optimal production and maintenance planning of biopharmaceutical manufacturing under performance decay. *Industrial and Engineering Chemistry Research*, 53(44):17075–17091, 11 2014. ISSN 15205045. doi: 10.1021/ie5008807.

S. Liu, L. G. Papageorgiou, and N. Shah. Optimal design of low-cost supply chain networks on the benefits of new product formulations. *Computers and Industrial Engineering*, 139(106189):1–17, 1 2020. ISSN 03608352. doi: 10.1016/j.cie.2019.106189.

M. Marani, G. G. Katul, W. K. Pan, and A. J. Parolari. Intensity and frequency of extreme novel epidemics. *PNAS Environmental sciences Biological sciences*, 118(35):1–4, 2021. doi: 10.1073/pnas.2105482118/-/DCSupplemental. URL `https://ghsagenda.org/`.

C. T. Maravelias and I. E. Grossmann. Simultaneous planning for new product development and batch manufacturing facilities. *Industrial and Engineering Chemistry Research*, 40(26):6147–6164, 2001. ISSN 08885885. doi: 10.1021/ie010301x.

C. T. Maravelias and I. E. Grossmann. A hybrid MILP/CP decomposition approach for the continuous time scheduling of multipurpose batch plants. *Computers and Chemical Engineering*, 28(10):1921–1949, 9 2004. ISSN 00981354. doi: 10.1016/j.compchemeng.2004.03.016.

C. M. Marques, S. Moniz, J. P. de Sousa, and A. P. Barbosa-Póvoa. A simulation-optimization approach to integrate process design and planning decisions under technical and market uncertainties: A case from the chemical-pharmaceutical industry. *Computers and Chemical Engineering*, 106:796–813, 11 2017. ISSN 00981354. doi: 10.1016/j.compchemeng.2017.04.008.

C. M. Marques, S. Moniz, J. P. de Sousa, A. P. Barbosa-Povoa, and G. Reklaitis. Decision-support challenges in the chemical-pharmaceutical industry: Findings and future research directions, 3 2020. ISSN 00981354.

O.-W. Merten, S. Charrier, N. Laroudie, S. Fauchille, C. Dugué, C. Jenny, M. Audit, M.-A. Zanta-Boussif, H. Chautard, M. Radrizzani, G. Vallanti, L. Naldini, P. Noguiez-Hellin, and A. Galy. Large-Scale Manufacture and Characterization of a Lentiviral Vector Produced for Clinical Ex Vivo Gene Therapy Application. *Human Gene Therapy*, 22(3):343–356, 11 2011. ISSN 1043-0342. doi: 10.1089/hum.2010.060. URL `https://doi.org/10.1089/hum.2010.060`.

O.-W. Merten, M. Schweizer, P. Chahal, and A. A. Kamen. Manufacturing of viral vectors for gene therapy: part I. Upstream processing. *Pharmaceutical Bioprocessing*, 2(2):183–203, 4 2014. ISSN 2048-9145. doi: 10.4155/pbp.14.16.

A. M. Mestre, M. D. Oliveira, and A. P. Barbosa-Póvoa. Location-allocation approaches for hospital network planning under uncertainty. *European Journal of Operational Research*, 240(3):791–806, 2015. ISSN 03772217. doi: 10.1016/j.ejor.2014.07.024.

M. G. Moleirinho, R. J. Silva, P. M. Alves, M. J. Carrondo, and C. Peixoto. Current challenges in biotherapeutic particles manufacturing. *Expert Opinion on Biological Therapy*, 2019. ISSN 17447682. doi: 10.1080/14712598.2020.1693541.

S. Moniz, A. Paula Barbosa-Povoa, and J. Pinho De Sousa. Scheduling with equipment redesign in multipurpose batch plants. 2012. URL `https://www.researchgate.net/publication/272784668`.

S. Moniz, A. P. Barbosa-Póvoa, and J. P. de Sousa. Simultaneous regular and non-regular production scheduling of multipurpose batch plants: A real chemical-pharmaceutical case study. *Computers and Chemical Engineering*, 67:83–102, 8 2014. ISSN 00981354. doi: 10.1016/j.compchemeng.2014.03.017.

C. P. Neves, J. L. Coffman, and S. S. Farid. Evaluating end-to-end continuous antibody manufacture with column-free capture alternatives from economic, environmental, and robustness perspectives. *Biotechnology Progress*, 2024. ISSN 15206033. doi: 10.1002/btpr.3427.

F. B. Oyebolu, J. van Lidth de Jeude, C. Siganporia, S. S. Farid, R. Allmendinger, and J. Branke. A new lot sizing and scheduling heuristic for multi-site biopharmaceutical production. *Journal of Heuristics*, 23(4):231–256, 8 2017. ISSN 15729397. doi: 10.1007/s10732-017-9338-9.

F. B. Oyebolu, R. Allmendinger, S. S. Farid, and J. Branke. Dynamic scheduling of multi-product continuous biopharmaceutical facilities: A hyper-heuristic framework. *Computers and Chemical Engineering*, 125:71–88, 6 2019. ISSN 00981354. doi: 10.1016/j.compchemeng.2019.03.002.

A. Panteli, S. Giarola, and N. Shah. Supply Chain Mixed Integer Linear Program Model Integrating a Biorefining Technology Superstructure. *Industrial and Engineering Chemistry Research*, 57(30):9849–9865, 8 2018. ISSN 15205045. doi: 10.1021/acs.iecr.7b05228.

C. C. Pantelides. Unified frameworks for optimal process planning and scheduling. *2nd Conference on Foundations of Computer-Aided Operations*, pages 253–274, 1994.

L. G. Papageorgiou and C. C. Pantelides. Optimal Campaign Planning/Scheduling of Multipurpose Batch/ Semicontinuous Plants. 2. A Mathematical Decomposition Approach. *Industrial Engineering and Chemistry Research*, 35: 510–529, 1996. URL `https://pubs.acs.org/sharingguidelines`.

L. G. Papageorgiou, G. E. Rotstein, and N. Shah. Strategic supply chain optimization for the pharmaceutical industries. *Industrial and Engineering Chemistry Research*, 40(1):275–286, 2001. ISSN 08885885. doi: 10.1021/ie990870t.

M. M. Papathanasiou, C. Stamatis, M. Lakelin, S. Farid, N. Titchener-Hooker, and N. Shah. Autologous CAR T-cell therapies supply chain: challenges and opportunities? *Cancer Gene Therapy*, 2020. ISSN 14765500. doi: 10.1038/s41417-019-0157-z.

E. Perea-López, I. E. Grossmann, B. E. Ydstie, and T. Tahmassebi. Dynamic modeling and decentralized control of supply chains. *Industrial and Engineering Chemistry Research*, 40(15):3369–3383, 7 2001. ISSN 08885885. doi: 10.1021/ie000573k.

E. Perea-López, B. E. Ydstie, and I. E. Grossmann. A model predictive control strategy for supply chain optimization. *Computers and Chemical Engineering*, 27(8-9):1201–1218, 9 2003. ISSN 00981354. doi: 10.1016/S0098-1354(03)00047-4.

T. D. Pereira Chilima, F. Moncaubeig, and S. S. Farid. Impact of allogeneic stem cell manufacturing decisions on cost of goods, process robustness and reimbursement. *Biochemical Engineering Journal*, 137:132–151, 9 2018. ISSN 1873295X. doi: 10.1016/j.bej.2018.04.017.

T. D. Pereira Chilima, F. Moncaubeig, and S. S. Farid. Estimating capital investment and facility footprint in cell therapy facilities. *Biochemical Engineering Journal*, 155, 3 2020. ISSN 1873295X. doi: 10.1016/j.bej.2019.107439.

C. Perry and A. C. M. E. Rayat. Lentiviral Vector Bioprocessing. *Viruses*, 13 (2):268, 2 2021. ISSN 1999-4915. doi: 10.3390/v13020268.

Peter Jones. Lentiviral vector CMC considerations for clinical and commercial use amc 19th Technical Meeting: Manufacturing ATMPs at Scale Dublin. Technical report, 2019.

Pfitzer. Net Zero by 2040: How Pfizer is Fighting Climate Change with Ambitious Science Based Goals, 2022. URL `https://www.pfizer.com/news/articles/net_zero_by_2040_how_pfizer_is_fighting_climate_change_with_ambitious_science_based_goals`.

Pharmaceutical Technology. Pharma's path to Net Zero: Targeting Scope 3 emissions , 2023. URL `https://www.pharmaceutical-technology.com/features/pharmas-path-to-net-zero-targeting-scope-3-emissions/`.

M. Pietrzykowski, W. Flanagan, V. Pizzi, A. Brown, A. Sinclair, and M. Monge. An environmental life cycle assessment comparison of single-use and conventional process technology for the production of monoclonal antibodies. *Journal of Cleaner Production*, 41:150–162, 2013. ISSN 09596526. doi: 10.1016/j.jclepro.2012.09.048.

J. M. Pinto and I. E. Grossmann. A Continuous Time Mixed Integer Linear Programming Model for Short Term Scheduling of Multistage Batch Plants The. Technical report, 1995. URL `https://pubs.acs.org/sharingguidelines`.

W. Qu, M. Wang, Y. Wu, and R. Xu. Scalable Downstream Strategies for Purification of Recombinant Adeno-Associated Virus Vectors in Light of the Properties. *Current Pharmaceutical Biotechnology*, 16:684–695, 2015.

G. Reklaitis. Perspectives on Scheduling and Planning of Process Operations. *4th International Symposium on Process Systems Engineering*, 1991.

G. E. Rotstein, L. G. Papageorgiou, N. Shah, D. C. Murphy, and R. Mustafa. A Product Portfolio Approach in the Pharmaceutical Industry. *Computers and Chemical Engineering*, 23(SUPPL. 1), 6 1999. ISSN 00981354. doi: 10.1016/S0098-1354(99)80217-8.

J. H. Ryu and E. N. Pistikopoulos. Multiperiod planning of enterprise-wide supply chains using an operation policy. *Industrial and Engineering Chemistry Research*, 46(24):8058–8065, 11 2007. ISSN 08885885. doi: 10.1021/ie070508b.

J. H. Ryu, V. Dua, and E. N. Pistikopoulos. Proactive scheduling under uncertainty: A parametric optimization approach. *Industrial and Engineering Chemistry Research*, 46(24):8044–8049, 11 2007. ISSN 08885885. doi: 10.1021/ie070018j.

N. Sabio, A. Kostin, G. Guillén-Gosálbez, and L. Jiménez. Holistic minimization of the life cycle environmental impact of hydrogen infrastructures using multi-objective optimization and principal component analysis. *International Journal of Hydrogen Energy*, 37(6):5385–5405, 3 2012. ISSN 03603199. doi: 10.1016/j.ijhydene.2011.09.039.

M. I. G. Salema, A. P. Barbosa-Povoa, and A. Q. Novais. An optimization model for the design of a capacitated multi-product reverse logistics network with uncertainty. *European Journal of Operational Research*, 179(3):1063–1077, 6 2007. ISSN 03772217. doi: 10.1016/j.ejor.2005.05.032.

J. J. Samaras, M. Micheletti, and W. Ding. Transformation of Biopharmaceutical Manufacturing Through Single-Use Technologies: Current State, Remaining Challenges, and Future Development. 2022. doi: 10.1146/annurev-chembioeng. URL `https://doi.org/10.1146/annurev-chembioeng-`.

T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115, 11 2005. ISSN 03772217. doi: 10.1016/j.ejor.2004.01.046.

Sartorius. Operating Instructions Original Operating Instructions Flexsafe STR ® Single-Use Bags Volumes 50 L-2000 L Operating Instructions Flexsafe STR ® Single-Use Bags 3 Contents Contents. Technical report, a.

Sartorius. Product Datasheet Product Information Features and Benefits-Accelerate Go-to-Market timelines-Simplify process transfer from PD to CM-Flexibility for different production scenarios and testing multiple setups-Precise process control for better parameter definition-Excellent global application support Biostat® RM Rocking Motion Bioreactor Powered by Biobrain®. Technical report, b.

G. Schilling and C. C. Pantelides. A simple continuous-time scheduling formulation and a novel solution algorithm. *Computers chem. Engng*, 20:1221–1226, 1996.

G. Schilling and C. C. Pantelides. Optimal Periodic Scheduling of Multipurpose Plants in the Continuous Time Domain. *Computers and Chemical Engineering*, 21:S1191–S1196, 1997.

C. W. Schmidt and I. E. Grossmann. Optimization Models for the Scheduling of Testing Tasks in New Product Development. *Industry & Chemical Engineering*, 35:3498–3510, 1996. URL `https://pubs.acs.org/sharingguidelines`.

N. Shah. Single- and multisite planning and scheduling: current status and future challenges. *3rd International Conference on Foundations of Computer-Aided Process Operations*, pages 75–90, 1998.

N. Shah, C. C. Pantelides, and R. W. H. Sargent. A General Algorithm for Short-Term Scheduling of Batch Operations II. Computational Issues. *Computers & Chemical Engineering*, 17(2):229–244, 1993.

C. Sheridan. Why gene therapies must go virus free. *Nature Biotechnology*, 41: 737–744, 2023.

C. C. Siganporia, S. Ghosh, T. Daszkowski, L. G. Papageorgiou, and S. S. Farid. Capacity planning for batch and perfusion bioprocesses across multiple biopharmaceutical facilities. *Biotechnology Progress*, 30(3):594–606, 2014. ISSN 15206033. doi: 10.1002/btpr.1860.

R. Sousa, N. Shah, and L. G. Papageorgiou. Supply chain design and multilevel planning-An industrial case. *Computers and Chemical Engineering*, 32(11):2643–2663, 11 2008. ISSN 00981354. doi: 10.1016/j.compchemeng.2007.09.005.

R. T. Sousa, S. Liu, L. G. Papageorgiou, and N. Shah. Global supply chain planning for pharmaceuticals. *Chemical Engineering Research and Design*, 89 (11):2396–2409, 11 2011. ISSN 02638762. doi: 10.1016/j.cherd.2011.04.005.

St. Gobain. Bioprocess Bags For single-use bioprocess systems. Technical report.

H. Stefansson, N. Shah, and P. Jensson. Multiscale planning and scheduling in the secondary pharmaceutical industry. *AIChE Journal*, 52(12):4133–4149, 2006. ISSN 15475905. doi: 10.1002/aic.10989.

H. Stefansson, S. Sigmarsdottir, P. Jensson, and N. Shah. Discrete and continuous time representations and mathematical models for large production scheduling problems: A case study from the pharmaceutical industry. *European Journal of Operational Research*, 215(2):383–392, 12 2011. ISSN 03772217. doi: 10.1016/j.ejor.2011.06.021.

D. Subramanian, J. F. Pekny, and G. V. Reklaitis. A Simulation-Optimization Framework for Research and Development Pipeline Management. *AIChE*, 47 (10):2226–2242, 2001.

K. Subramanian, J. B. Rawlings, C. T. Maravelias, J. Flores-Cerrillo, and L. Megan. Integration of control theory and scheduling methods for supply chain management. *Computers and Chemical Engineering*, 51:4–20, 4 2013. ISSN 00981354. doi: 10.1016/j.compchemeng.2012.06.012.

A. Sundaramoorthy and C. T. Maravelias. Simultaneous batching and scheduling in multistage multiproduct processes. *Industrial and Engineering Chemistry Research*, 47(5):1546–1555, 3 2008. ISSN 08885885. doi: 10.1021/ie070944y.

A. Sundaramoorthy, J. M. Evans, and P. I. Barton. Capacity planning under clinical trials uncertainty in continuous pharmaceutical manufacturing, 1: Mathematical framework. *Industrial and Engineering Chemistry Research*, 51 (42):13692–13702, 10 2012a. ISSN 08885885. doi: 10.1021/ie300324h.

A. Sundaramoorthy, X. Li, J. M. Evans, and P. I. Barton. Capacity planning under clinical trials uncertainty in continuous pharmaceutical manufacturing, 2: Solution method. *Industrial and Engineering Chemistry Research*, 51(42): 13703–13711, 10 2012b. ISSN 08885885. doi: 10.1021/ie3003254.

N. Susarla and I. A. Karimi. Integrated supply chain planning for multinational pharmaceutical enterprises. *Computers and Chemical Engineering*, 42:168–177, 7 2012. ISSN 00981354. doi: 10.1016/j.compchemeng.2012.03.002.

B. A. Thorne, R. K. Takeya, and R. W. Peluso. Manufacturing Recombinant Adeno-Associated Viral Vectors from Producer Cell Clones. *Human Gene Therapy*, 20:707–714, 2009. URL `www.liebertpub.com`.

N. Triantafyllou and M. M. Papathanasiou. Deep learning enhanced mixed integer optimization: Learning to reduce model dimensionality. *Computers and Chemical Engineering*, 187, 8 2024. ISSN 00981354. doi: 10.1016/j.compchemeng.2024.108725.

N. Triantafyllou, A. Bernardi, M. Lakelin, N. Shah, and M. M. Papathanasiou. A digital platform for the design of patient-centric supply chains. *Scientific Reports*, 12(1), 12 2022. ISSN 20452322. doi: 10.1038/s41598-022-21290-5.

K. H. Tsang, N. J. Samsatli, and N. Shah. Capacity investment planning for multiple vaccines under uncertainty: 1: Capacity planning. *Food and Bioproducts Processing*, 85(2 C):120–128, 2007a. ISSN 09603085. doi: 10.1205/fbp06001.

K. H. Tsang, N. J. Samsatli, and N. Shah. Capacity investment planning for multiple vaccines under uncertainty: 2: Financial risk analysis. *Food and Bioproducts Processing*, 85(2 C):129–140, 2007b. ISSN 09603085. doi: 10.1205/fbp06002.

P. Tsiakis and L. G. Papageorgiou. Optimal production allocation and distribution supply chain networks. *International Journal of Production Economics*, 111(2):468–483, 2 2008. ISSN 09255273. doi: 10.1016/j.ijpe.2007.02.035.

P. Tsiakis, N. Shah, and C. C. Pantelides. Design of multi-echelon supply chain networks under demand uncertainty. *Industrial and Engineering Chemistry Research*, 40(16):3585–3604, 8 2001. ISSN 08885885. doi: 10.1021/ie0100030.

U.S. Food & Drug Administration (FDA). FDA approves novel gene therapy to treat patients with a rare form of inherited vision loss, 2017. URL https://www.fda.gov/news-events/press-announcements/fda-approves-novel-gene-therapy-treat-patients-rare-form-inherited-vision-loss.

D. Van Riel and E. Wit. Next-generation vaccine platforms for COVID-19. *Nature Materials*, 2020. doi: 10.1038/s41563-020-0746-0. URL https://doi.org/10.1038/s41563-020-0746-0.

J. Vellinga, J. P. Smith, A. Lipiec, D. Majhen, A. Lemckert, M. Van Ooij, P. Ives, C. Yallop, J. Custers, and M. Havenga. Challenges in manufacturing adenoviral vectors for global vaccine product deployment. *Human Gene Therapy*, 25 (4):318–327, 4 2014. ISSN 15577422. doi: 10.1089/hum.2014.007.

M. Vieira, T. Pinto-Varela, S. Moniz, A. P. Barbosa-Póvoa, and L. G. Papageorgiou. Optimal planning and campaign scheduling of biopharmaceutical processes using a continuous-time formulation. *Computers and Chemical Engineering*, 91:422–444, 9 2016. ISSN 00981354. doi: 10.1016/j.compchemeng.2016.04.009.

M. Vieira, T. Pinto-Varela, and A. P. Barbosa-Póvoa. Production and maintenance planning optimisation in biopharmaceutical processes under performance decay using a continuous-time formulation: A multi-objective approach. *Computers and Chemical Engineering*, 107:111–139, 12 2017. ISSN 00981354. doi: 10.1016/j.compchemeng.2017.03.008.

M. Vieira, T. Pinto-Varela, and A. P. Barbosa-Póvoa. A model-based decision support framework for the optimisation of production planning in the biopharmaceutical industry. *Computers and Industrial Engineering*, 129:354–367, 3 2019. ISSN 03608352. doi: 10.1016/j.cie.2019.01.045.

J. Whitley, C. Zwolinski, C. Denis, M. Maughan, L. Hayles, D. Clarke, M. Snare, H. Liao, S. Chiou, T. Marmura, H. Zoeller, B. Hudson, J. Peart, M. Johnson, A. Karlsson, Y. Wang, C. Nagle, C. Harris, D. Tonkin, S. Fraser, L. Capiz, C. L. Zeno, Y. Meli, D. Martik, D. A. Ozaki, A. Caparoni, J. E. Dickens, D. Weissman, K. O. Saunders, B. F. Haynes, G. D. Sempowski, T. N. Denny, and M. R. Johnson. Development of mRNA manufacturing for vaccines and

therapeutics: mRNA platform requirements and development of a scalable production process to support early phase clinical trials, 4 2022. ISSN 18781810.

WHO. Total COVID-19 vaccine doses administered, 2024.

J. Wiebe, I. CecÍlio, and R. Misener. Robust Optimization for the Pooling Problem. *Industrial and Engineering Chemistry Research*, 58(28):12712–12722, 7 2019. ISSN 15205045. doi: 10.1021/acs.iecr.9b01772.

C. H. Wong, D. Li, N. Wang, J. Gruber, A. W. Lo, and R. M. Conti. The estimated annual financial impact of gene therapy in the United States. *Gene Therapy*, 30(10-11):761–773, 11 2023. ISSN 14765462. doi: 10.1038/s41434-023-00419-9.

D. Wu and M. Ierapetritou. Hierarchical approach for production planning and scheduling under uncertainty. *Chemical Engineering and Processing: Process Intensification*, 46(11):1129–1140, 11 2007. ISSN 02552701. doi: 10.1016/j.cep.2007.02.021.

O. Yang, M. Qadan, and M. Ierapetritou. Economic Analysis of Batch and Continuous Biopharmaceutical Antibody Production: a Review, 3 2020. ISSN 19398042.

H. Yin, R. L. Kanasty, A. A. Eltoukhy, A. J. Vegas, J. R. Dorkin, and D. G. Anderson. Non-viral vectors for gene-based therapy, 2014. ISSN 14710064.

J. C. Zapata, J. Pekny, and G. V. Reklaitis. Simulation-Optimization in Support of Tactical and Strategic Enterprise Decisions. In K. G. Kempf, P. Keskinocak, and R. Uzsoy, editors, *Planning Production and Inventories in the Extended Enterprise: A State of the Art Handbook, Volume 1*, pages 593–627. Springer US, New York, NY, 2011. ISBN 978-1-4419-6485-4. doi: 10.1007/978-1-4419-6485-4{_}20. URL `https://doi.org/10.1007/978-1-4419-6485-4_20`.

L. J. Zeballos, C. A. Méndez, A. P. Barbosa-Povoa, and A. Q. Novais. Multi-period design and planning of closed-loop supply chains with uncertain supply and demand. *Computers and Chemical Engineering*, 66:151–164, 7 2014. ISSN 00981354. doi: 10.1016/j.compchemeng.2014.02.027.

Appendix A

Techno-economic modelling

A.1 Materials

A.1.1 AdV manufacturing

Table A.1 displays material inputs for the AdV process used in Chapter 3. The USP part of the process uses single-use equipment, where the DSP part use stainless steel equipment. Hence, material flows for phosphoric acid, sodium hydroxide and water for SIP/CIP are non-zero.

Table A.1: AdV mixed manufacturing inputs (kg/batch)

Material Input	2000	1000	200	50
AdV	0	0	0	0
Air	2,954.49	1,477.24	295.45	73.86
Benzonase	0.01	0	0	0
Biomass	0	0	0	0
Box	105.28	52.64	10.53	2.63
Caps	105.28	52.64	10.53	2.63
Carb. Dioxide	14.27	7.13	1.43	0.36
Elution Buffer	81.58	40.79	8.16	2.04
Formulation Buf	8,717.15	4,358.57	871.71	217.93
H3PO4 (5% w/w)	9,695.40	7,695.24	4,500.21	2,834.95
Labels	105.28	52.64	10.53	2.63
Loading Buffer	8,301.75	4,150.88	830.18	207.54
Lysis Buffer	532.35	266.17	53.23	13.31
NaOH (0.5 M)	4,713.81	3,741.35	2,187.96	1,378.33
O2	154.37	77.19	15.44	3.91
PBS	1,000.00	500	100	25
Regen Buffer	1,716.66	858.33	171.67	42.92
Serum Free Medi	1,985.43	992.72	198.54	49.64
Vials	523.62	261.81	52.36	13.09
Water	8,732.32	6,126.64	3,058.26	1,824.13
WFI	18,487.31	14,648.77	8,556.10	5,393.65

Table A.2 displays material inputs for the AdV process used in Chapters 6 and 7. The USP, DSP and FF part of the process use single-use equipment. Hence, material flows for phosphoric acid, sodium hydroxide and water for SIP/CIP are zero.

Table A.2: AdV SU manufacturing inputs (kg/batch)

Material Input	2000	1000	200	50
AdV	0	0	0	0
Air	2,954.49	1,477.24	295.45	73.86
Benzonase	0.01	0	0	0
Biomass	0	0	0	0
Box	105.28	52.64	10.53	2.63
Caps	105.28	52.64	10.53	2.63
Carb. Dioxide	14.27	7.13	1.43	0.36
Elution Buffer	81.58	40.79	8.16	2.04
Formulation Buf	8,717.15	400.24	80.05	20.01
Labels	105.28	52.64	10.53	2.63
Loading Buffer	8,301.75	1,734.99	347	86.75
Lysis Buffer	532.35	266.17	53.23	13.31
O2	154.37	77.19	15.44	3.91
PBS	1,000.00	500	100	25
Regen Buffer	1,716.66	306	61.2	15.3
Serum Free Medi	1,985.43	992.72	198.54	49.64
Vials	523.62	261.81	52.36	13.09
WFI	117.68	68.79	29.67	22.34

Table A.2 displays material inputs for the AdV process used in Chapter 3. The USP, DSP and FF part of the process use stainless steel equipment. Hence, material flows for phosphoric acid, sodium hydroxide and water for SIP/CIP are non-zero.

Table A.3: AdV MU manufacturing inputs (kg/batch)

Material Input	2000	1000	200	50
AdV	0.00	0.00	0.00	0.00
Air	2,954.49	1,477.24	295.45	73.86
Benzonase	0.01	0.00	0.00	0.00
Biomass	0.00	0.00	0.00	0.00
Box	105.28	52.64	10.53	2.63
Caps	105.28	52.64	10.53	2.63
Carb. Dioxide	14.27	7.13	1.43	0.36
Elution Buffer	81.58	40.79	8.16	2.04
Formulation Buf	8,717.15	4,358.57	871.71	217.93
H3PO4 (5% w/w)	9,695.40	7,695.24	4,500.21	2,834.95
Labels	105.28	52.64	10.53	2.63
Loading Buffer	8,301.75	4,150.88	830.18	207.54
Lysis Buffer	532.35	266.17	53.23	13.31
NaOH (0.5 M)	4,713.81	3,741.35	2,187.96	1,378.33
O2	154.37	77.19	15.44	3.91
PBS	1,000.00	500.00	100.00	25.00
Regen Buffer	1,716.66	858.33	171.67	42.92
Serum Free Medi	1,985.43	992.72	198.54	49.64
Vials	523.62	261.81	52.36	13.09
Water	8,732.32	6,126.64	3,058.26	1,824.13
WFI	18,487.31	14,648.77	8,556.10	5,393.65

A.1.2 LV manufacturing

Table A.4 displays material inputs for the LV process used in Chapters 6 and 7. The USP part of the process uses single-use equipment, where the DSP part use stainless steel equipment. Hence, material flows for phosphoric acid, sodium hydroxide and water for SIP/CIP are non-zero.

Table A.4: LV mixed manufacturing inputs (kg/batch) by scale

Material Input	2000	1000	200	50
Air	2,954.49	1,477.24	295.45	73.86
Benzonase	0.00	0.00	0.00	0.00
Biomass	0.00	0.00	0.00	0.00
Box	0.11	0.06	0.01	0.00
Caps	0.11	0.06	0.01	0.00
Carb. Dioxide	14.27	7.13	1.43	0.36
Elution Buffer	81.52	40.76	8.15	2.04
Formulation Buf	8,710.96	4,355.48	871.10	217.77
H3PO4 (5% w/w)	7,486.09	5,941.72	3,474.74	2,188.95
Labels	0.11	0.06	0.01	0.00
Loading Buffer	6,317.27	3,158.63	631.73	157.93
NaOH (0.5 M)	3,639.67	2,888.81	1,689.38	1,064.25
O2	154.37	77.19	15.44	3.91
PBS	10.00	5.00	1.00	0.25
pDNA	0.00	0.00	0.00	0.00
Regen Buffer	1,715.44	857.72	171.54	42.89
Serum Free Medi	1,984.03	992.01	198.40	49.60
Vials	5.67	2.83	0.57	0.14
Water	5,724.95	3,875.73	1,829.63	1,068.33
WFI	14,209.35	11,275.41	6,595.01	4,157.68

Table A.5 displays material inputs for the LV process used in Chapters 6 and 7. The USP, DSP and FF part of the process use single-use equipment. Hence, material flows for phosphoric acid, sodium hydroxide and water for SIP/CIP are zero.

Table A.5: LV SU manufacturing inputs (kg/batch) by scale

Material Input	2000	1000	200	50
Air	2,954.49	1,477.24	295.45	73.86
Benzonase	0.00	0.00	0.00	0.00
Biomass	0.00	0.00	0.00	0.00
Box	0.11	0.06	0.01	0.00
Caps	0.11	0.06	0.01	0.00
Carb. Dioxide	14.27	7.13	1.43	0.36
Elution Buffer	81.52	40.76	8.15	2.04
Formulation Buf	20.01	10.01	2.00	0.50
Labels	0.11	0.06	0.01	0.00
Loading Buffer	170.28	85.14	17.03	4.26
O2	154.37	77.19	15.44	3.91
PBS	10.00	5.00	1.00	0.25
pDNA	0.00	0.00	0.00	0.00
Regen Buffer	102.00	51.00	10.20	5.10
Serum Free Medi	1,984.03	992.01	198.40	49.60
Vials	5.67	2.83	0.57	0.14
WFI	25.63	17.79	11.51	10.34

Table A.5 displays material inputs for the LV process used in Chapters 6 and 7. The USP, DSP and FF part of the process use stainless steel equipment. Hence, material flows for phosphoric acid, sodium hydroxide and water for SIP/CIP are non-zero.

Table A.6: LV MU manufacturing inputs (kg/batch) by scale

Material	2000	1000	200	50
Air	2,954.48	1,477.24	295.45	73.86
Benzonase	0.00	0.00	0.00	0.00
Biomass	0.00	0.00	0.00	0.00
Box	0.11	0.06	0.01	0.00
Caps	0.11	0.06	0.01	0.00
Carb. Dioxide	14.27	7.13	1.43	0.36
Elution Buffer	81.52	40.76	8.15	2.04
Formulation Buf	20.01	10.01	2.00	0.50
H3PO4 (5% w/w)	6,685.48	2,078.61	1,215.58	803.95
Labels	0.11	0.06	0.01	0.00
Loading Buffer	170.28	85.14	17.03	4.26
NaOH (0.5 M)	4,770.81	1,010.60	591.00	390.87
O2	154.37	77.19	15.44	3.91
PBS	10.00	5.00	1.00	0.25
pDNA	0.00	0.00	0.00	0.00
Regen Buffer	102.00	51.00	10.20	5.10
Serum Free Medi	1,984.03	992.01	198.40	49.60
Vials	5.67	2.83	0.57	0.14
Water	4,677.16	10,058.48	5,755.76	3,603.22
WFI	11,905.08	3,956.08	2,314.64	1,533.56

A.1.3 AAV manufacturing

Table A.7 displays material inputs for the AAV process used in Chapter 3. The USP part of the process uses single-use equipment, where the DSP part use stainless steel equipment. Hence, material flows for phosphoric acid, sodium hydroxide and water for SIP/CIP are non-zero.

Table A.7: AAV mixed manufacturing inputs (kg/batch) by scale

Material Input	2000	1000	200	50	
Air	75763.92	35578.98	7115.80	1778.95	kg/batch
Benzonase	0.01	0.00	0.00	0.00	kg/batch
Biomass	0.00	0.00	0.00	0.00	kg/batch
Box	1.61	0.76	0.15	0.04	kg/batch
Caps	1.61	0.76	0.15	0.04	kg/batch
Carb. Dioxide	168.92	91.67	18.33	4.58	kg/batch
Elution Buffer	163.15	81.58	16.32	4.08	kg/batch
Formulation Buf	8717.15	4358.57	871.71	217.93	kg/batch
H3PO4 (5% w/w)	10814.59	8583.55	5019.69	3162.21	kg/batch
Labels	1.61	0.76	0.15	0.04	kg/batch
Loading Buffer	14623.50	7311.75	1462.35	365.59	kg/batch
Lysis Buffer	532.35	266.17	53.23	13.31	kg/batch
NaOH (0.5 M)	5257.95	4173.24	2440.52	1537.43	kg/batch
O2	273.68	161.69	32.34	8.08	kg/batch
PBS	10.00	5.00	1.00	0.25	kg/batch
pDNA	0.00	0.00	0.00	0.00	kg/batch
Regen Buffer	3433.32	1716.66	343.33	85.83	kg/batch
Serum Free Medi	1985.43	992.72	198.54	49.64	kg/batch
Vials	80.31	37.74	7.55	1.89	kg/batch
Water	9549.76	6505.21	3109.17	1822.26	kg/batch
WFI	20629.28	16339.10	9541.81	6014.05	kg/batch

A.2 Mixed equipment flowsheets

A.2.1 AdV manufacturing

Table A.8 displays equipment list for the AdV process used in Chapter 3. The USP part of the process uses single-use equipment, where the DSP part use stainless steel equipment. Hence, skids to support disposable bags are installed in USP and tanks are installed in DSP, requiring SIP/CIP steps.

Table A.8: AdV mixed equipment summary by step

		2000		1000		200		50	
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
1	Media Hold Skid	1	100	1	100	1	100	1	100	L
1	Media Prep Skid	1	100	1	100	1	100	1	100	L
1	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
1	Rocking Bioreactor Skid	1	50.00	1	50.00	1	50.00	1	50.00	L
1	Test Tube Rack	1	0.01	1	0.01	1	0.01	1	0.01	L
2	Media Prep Skid	1	100	1	100	1	100	1	100	L
2	Media Hold Skid	1	100	1	100	1	100	1	100	L
2	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
2	Rocking Bioreactor Skid	1	200.00	1	200.00	1	200.00	1	200.00	L
3	Air Filter	1	1,444.76	1	722.38	1	144.48	1	36.12	L/h
3	Media Hold Skid	1	1000	1	1000	1	1000	1	100	L
3	Medium Hold Skid	1	1000	1	1000	1	1000	1	100	L
3	Centrifugal Compressor	1	0.76	1	0.38	1	0.08	1	0.02	kW
3	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
3	Mixer	1	9.69	1	4.84	1	0.97	1	0.24	kg/h
3	Seed Bioreactor	1	349.71	1	174.85	1	34.97	1	8.74	L
4	Air Filter	1	11,792.49	1	5,896.26	1	1,179.25	1	294.81	L/h
4	Bioreactor	1	2,502.46	1	1,251.23	1	250.25	1	62.56	L
4	Media Hold Tank	1	1,905.54	1	952.77	1	190.55	1	47.64	L
4	Media Prep Tank	1	1,905.54	1	952.77	1	190.55	1	47.64	L
4	Centrifugal Compressor	1	6.17	1	3.09	1	0.62	1	0.15	kW
4	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
4	Disposable Generic Container Skid	1	58.00	1	29.00	1	6.00	1	2.00	L
4	Mixer	1	79.06	1	39.53	1	7.91	1	1.98	kg/h
4	Roller Bottle Rack	1	2.20	1	2.20	1	2.20	1	2.20	L
5	Lysis Tank	1	2,798.92	1	1,399.46	1	279.89	1	69.97	L
5	Lysis Hold Tank	1	575.16	1	287.58	1	57.52	1	14.38	L
5	Lysis Buffer Prep Tank	1	575.16	1	287.58	1	57.52	1	14.38	L
5	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
6	Disk-Stack Centrifuge	1	627.57	1	313.78	1	62.76	1	15.69	L/h
7	Dead-End Filter	1	3.50	1	1.75	1	0.35	1	0.09	m2
8	Ultrafiltration Tank	1	2,798.46	1	1,402.36	1	283.50	1	71.62	L

Table A.8: AdV mixed equipment summary by step (ctn'd)

	2000		1000		200		50	
Type	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
Buffer Hold Tank	1	2,139.88	1	1,069.94	1	213.99	1	53.50	L
Buffer Prep Tank	1	2,139.88	1	1,069.94	1	213.99	1	53.50	L
Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
Diafilter	1	44.46	1	22.28	1	4.50	1	1.14	m2
Wash Buffer Prep Tank	1	2,277.41	1	1,138.70	1	227.74	1	56.94	L
Wash Buffer Hold Tank	1	2,277.41	1	1,138.70	1	227.74	1	56.94	L
Loading Buffer Hold Tank	1	4,554.81	1	2,277.41	1	455.48	1	113.87	L
Elution Hold Tank	1	114.70	1	57.35	1	11.47	1	2.87	L
Regen Buffer Tank	1	1,890.90	1	945.45	1	189.09	1	47.27	L
Regen Hold Tank	1	1,890.90	1	945.45	1	189.09	1	47.27	L
Elution Buffer Prep Tank	1	88.38	1	44.19	1	8.84	1	2.21	L
Elution Hold Tank	1	88.38	1	44.19	1	8.84	1	2.21	L
Loading Buffer Prep Tank	1	4,554.81	1	2,277.41	1	455.48	1	113.87	L
Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
IEX Column	1	7.65	1	3.83	1	0.77	1	0.19	L
Formulation Buffer Prep Tank	1	9,708.16	1	4,854.08	1	970.82	1	242.70	L
Buffer Hold Tank	1	9,708.16	1	4,854.08	1	970.82	1	242.70	L
Dead-End Filter	1	40.00	1	20.00	1	10.00	1	10.00	m2
Diafilter	1	3.08	1	1.54	1	0.31	1	0.08	m2
Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
Freeze-Thaw Module	1	57.36	1	28.68	1	5.74	1	1.43	L
Formulation Tank	1	1,169.53	1	584.77	1	116.95	1	29.24	L
Dead-End Filter	1	220.00	1	110.00	1	30.00	1	10.00	m2
Ultrafilter	1	9.33	1	4.66	1	0.93	1	0.23	m2
Assembler	5	350.94	3	292.45	1	175.47	1	43.87	entities/min
Discrete Freeze-Thaw Module	17	96.88	9	91.50	2	82.35	1	41.17	L
Filler	5	350.94	3	292.45	1	175.47	1	43.87	entities/min
Labeler	264	398.79	132	398.79	27	389.93	7	376.00	entities/h
Packer	11	382.84	6	350.94	2	210.56	1	105.28	entities/min

A.2.2 LV manufacturing

Table A.9 displays equipment list for the LV process used in Chapter 3. The USP part of the process uses single-use equipment, where the DSP part use stainless steel equipment. Hence, skids to support disposable bags are installed in USP and tanks are installed in DSP, requiring SIP/CIP steps.

Table A.9: LV mixed equipment summary by step

		2000		1000		200		50		
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
1	Air Filter	1	745.77	1	372.89	1	74.58	1	18.64	L/h
1	Medium Hold Skid	1	100	1	100	1	100	1	100	L
1	Medium Prep Skid	1	100	1	100	1	100	1	100	L
1	Centrifugal Compressor	1	0.39	1	0.2	1	0.04	1	0.01	kW
1	Dead-End Filter	1	10	1	10	1	10	1	10	m2
1	Mixer	1	5	1	2.5	1	0.5	1	0.13	kg/h
1	Rocking Bioreactor Skid	1	50	1	50	1	50	1	50	L
1	Test Tube Rack	1	0.01	1	0.01	1	0.01	1	0.01	L
2	Air Filter	1	496.32	1	248.16	1	49.63	1	12.41	L/h
2	Media Hold Skid	1	100	1	31.66	1	100	1	100	L
2	Media Prep Skid	1	100	1	31.66	1	100	1	100	L
2	Centrifugal Compressor	1	0.26	1	0.13	1	0.03	1	0.01	kW
2	Dead-End Filter	1	10	1	10	1	10	1	10	m2
2	Mixer	1	3.33	1	1.66	1	0.33	1	0.08	kg/h
2	Rocking Bioreactor Skid	1	200	1	200	1	200	1	200	L
3	Media Prep Skid	1	1000	1	1000	1	100	1	100	L
3	Media Hold Skid	1	1000	1	1000	1	100	1	100	L
3	Dead-End Filter	1	10	1	10	1	10	1	10	m2
3	Seed Bioreactor	1	349.46	1	174.73	1	34.95	1	8.74	L
4	Bioreactor	1	2,802.37	1	1,401.18	1	280.24	1	70.06	L
4	Media Hold Tank	1	1,904.19	1	952.09	1	190.42	1	47.61	L
4	Media Prep Tank	1	1,904.19	1	952.09	1	190.42	1	47.61	L
4	Dead-End Filter	1	10	1	10	1	10	1	10	m2
4	Roller Bottle Rack	1	2.2	1	2.2	1	2.2	1	2.2	L
5	Endonuclease Tank	1	2,112.93	1	1,056.96	1	212.18	1	53.8	L
6	Microfilter	1	2,223.10	1	1,111.55	1	222.31	1	55.58	L
6	Microfilter	1	36.19	1	36.19	1	36.19	1	36.19	m2
7	Wash Buffer Hold Tank	1	92.02	1	46.01	1	9.2	1	2.3	L

Table A.9: LV mixed equipment summary by step (ctn'd)

		2000		1000		200		50		
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
7	Wash Buffer Prep Tank	1	92.02	1	46.01	1	9.2	1	2.3	L
7	Regen Hold Tank	1	112.35	1	56.18	1	11.24	1	5.62	L
7	Regen Buffer Tank	1	112.35	1	56.18	1	11.24	1	5.62	L
7	Elution Buffer Tank	1	88.32	1	44.16	1	8.83	1	2.21	L
7	Loading Hold Tank	1	92.02	1	46.01	1	9.2	1	2.3	L
7	Loading Buffer Tank	1	92.02	1	46.01	1	9.2	1	2.3	L
7	Elution Hold Tank	1	88.32	1	44.16	1	8.83	1	2.21	L
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	IEX Column	1	1.18	1	0.59	1	0.12	1	0.03	L
8	Ultrafiltration Tank	1	9.19	1	4.6	1	0.92	1	0.23	L
8	Buffer Hold Tank	1	22.29	1	11.14	1	2.23	1	0.56	L
8	Formulation Tank	1	22.29	1	11.14	1	2.23	1	0.56	L
8	Dead-End Filter	1	10	1	10	1	10	1	10	m2
8	Diafilter	1	0.82	1	0.41	1	0.08	1	0.02	m2
8	Disposable Generic Container Skid	1	2	1	1	1	1	1	1	L
9	Dead-End Filter	1	10	1	10	1	10	1	10	m2
10	Dead-End Filter	1	10	1	10	1	10	1	10	m2
12	Storage Tank	1	12.65	1	6.33	1	1.27	1	0.32	L
12	Freeze-Thaw Module	1	1.53	1	0.77	1	0.15	1	0.04	L
12	Ultrafilter	1	0.25	1	0.12	1	0.03	1	0.01	m2
13	Assembler	1	113.93	1	56.96	1	11.39	1	2.85	entities/h
13	Filler	1	1.9	1	0.95	1	0.19	1	0.05	entities/min
13	Labeler	1	113.94	1	56.97	1	11.39	1	2.85	entities/h
13	Packer	1	4.56	1	2.28	1	0.46	1	0.11	entities/min
15	Discrete Freeze-Thaw Module	1	8.45	1	4.22	1	0.85	1	0.21	L

A.2.3 AAV manufacturing

Table A.10 displays equipment list for the AAV process used in Chapter 3. The USP part of the process uses single-use equipment, where the DSP part use stainless steel equipment. Hence, skids to support disposable bags are installed in USP and tanks are installed in DSP, requiring SIP/CIP steps.

Table A.10: AAV mixed equipment summary by step

		2000		1000		200		50		
Step	**Type**		**Capacity**		**Capacity**		**Capacity**		**Capacity**	
1	Air Filter	1	1,444.76	1	722.38	1	144.48	1	36.12	L/h
1	Centrifugal Compressor	1	0.76	1	0.38	1	0.08	1	0.02	kW
1	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
1	Mixer	1	9.69	1	4.84	1	0.97	1	0.24	kg/h
1	Rocking Bioreactor Skid	1	200.00	1	200.00	1	200.00	1	200.00	L
1	Roller Bottle Rack	1	2.20	1	2.20	1	2.20	1	2.20	L
1	Skid for Disposable Bag	1	100.00	1	100.00	1	100.00	1	100.00	L
1	Skid for Disposable Bag	1	100.00	1	100.00	1	100.00	1	100.00	L
1	Test Tube Rack	1	0.01	1	0.01	1	0.01	1	0.01	L
2	Air Filter	1	12.12	1	5.90	1	1.18	1	0.30	m3/h
2	Centrifugal Compressor	1	6.35	1	3.09	1	0.62	1	0.15	kW
2	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
2	Mixer	1	81.28	1	39.53	1	7.91	1	1.98	kg/h
2	Rocking Bioreactor Skid	1	50.00	1	50.00	1	50.00	1	50.00	L
2	Skid for Disposable Bag	1	100.00	1	100.00	1	100.00	1	100.00	L
2	Skid for Disposable Bag	1	100.00	1	100.00	1	100.00	1	100.00	L
3	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
3	Seed Bioreactor	1	349.71	1	174.85	1	34.97	1	8.74	L
3	Skid for Disposable Bag	1	500.00	1	500.00	1	500.00	1	500.00	L
3	Skid for Disposable Bag	1	500.00	1	500.00	1	500.00	1	500.00	L
4	Bioreactor	1	2,502.46	1	1,251.23	1	250.25	1	62.56	L
4	Medium Hold Tank	1	1,905.54	1	952.77	1	190.55	1	47.64	L
4	Meda Prep Tank	1	1,905.54	1	952.77	1	190.55	1	47.64	L
4	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
5	Lysis Tank	1	2,799.23	1	1,399.62	1	279.92	1	69.98	L
5	Lysis Hold Tank	1	575.16	1	287.58	1	57.52	1	14.38	L
5	Lysis Buffer Prep Tank	1	575.16	1	287.58	1	57.52	1	14.38	L
5	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
6	Microfilter	1	18.61	1	18.61	1	18.61	1	18.61	m2
7	Storage Tank	1	2,650.50	1	1,325.50	1	265.50	1	66.75	L
7	Buffer Hold Tank	1	2,139.88	1	1,069.94	1	213.99	1	53.50	L
7	Formulation Buffer Tank	1	2,139.88	1	1,069.94	1	213.99	1	53.50	L
7	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
7	Diafilter	1	19.36	1	15.84	1	3.17	1	0.80	m2
8	Wash Buffer Tank	1	2,277.41	1	1,138.70	1	227.74	1	56.94	L
8	Buffer Tank	1	106.69	1	106.69	1	106.69	1	106.69	L
8	Wash Buffer Tank	1	2,277.41	1	1,138.70	1	227.74	1	56.94	L
8	Regen Buffer Tank	1	1,890.90	1	945.45	1	189.09	1	47.27	L

Table A.10: AAV mixed equipment summary by step (cnt'd)

		2000		1000		200		50		
Step	**Type**		**Capacity**		**Capacity**		**Capacity**		**Capacity**	
8	Regen Hold Tank	1	1,890.90	1	945.45	1	189.09	1	47.27	L
8	Elution Buffer Prep Tank	1	88.38	1	44.19	1	8.84	1	2.21	L
8	Elution Hold Tank	1	88.38	1	44.19	1	8.84	1	2.21	L
8	Loading Buffer Prep Tank	1	4,554.81	1	2,277.41	1	455.48	1	113.87	L
8	Loading Hold Tank	1	4,554.81	1	2,277.41	1	455.48	1	113.87	L
8	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
8	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
8	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
8	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
8	IEX Column	1	2.58	1	1.29	1	0.26	1	0.06	L
9	Loading Buffer Prep Tank	1	4,554.81	1	2,277.41	1	455.48	1	113.87	L
9	Elution Buffer Prep Tank	1	88.38	1	44.19	1	8.84	1	2.21	L
9	Elution Hold Tank	1	88.38	1	44.19	1	8.84	1	2.21	L
9	Wash Buffer Tank	1	2,277.41	1	1,138.70	1	227.74	1	56.94	L
9	Wash Buffer Tank	1	2,277.41	1	1,138.70	1	227.74	1	56.94	L
9	Regen Buffer Tank	1	1,890.90	1	945.45	1	189.09	1	47.27	L
9	Regen Hold Tank	1	1,890.90	1	945.45	1	189.09	1	47.27	L
9	Loading Hold Tank	1	4,554.81	1	2,277.41	1	455.48	1	113.87	L
9	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
9	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
9	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
9	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
9	Affinity Column	1	6.75	1	3.11	1	0.62	1	0.16	L
10	Storage Tank	1	101.17	1	46.54	1	9.31	1	2.33	L
10	Buffer Hold Tank	1	9,708.16	1	4,854.08	1	970.82	1	242.70	L
10	Formulation Buffer Tank	1	9,708.16	1	4,854.08	1	970.82	1	242.70	L
10	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
10	Dead-End Filter	1	40.00	1	20.00	1	10.00	1	10.00	m2
10	Diafilter	1	3.96	1	1.82	1	0.36	1	0.09	m2
11	Disposable Container Skid	1	34.00	1	16.00	1	4.00	1	1.00	L
12	Freeze-Thaw Module	1	33.73	1	15.51	1	3.10	1	0.78	L
13	Formulation Tank	1	38.13	1	38.13	1	38.13	1	38.13	L
13	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
14	Ultrafilter	1	0.34	1	0.19	1	0.04	1	0.01	m2
15	Assembler	1	1,372.97	1	758.80	1	151.76	1	37.94	entities/h
15	Discrete Freeze-Thaw Module	2	50.90	1	56.27	1	11.25	1	2.81	L
15	Filler	1	22.88	1	12.65	1	2.53	1	0.63	entities/min
15	Labeler	1	1,372.97	1	758.80	1	151.76	1	37.94	entities/h
15	Packer	1	54.92	1	30.35	1	6.07	1	1.52	entities/min

A.3 MU equipment flowsheets

A.3.1 AdV manufacturing

Table A.11 displays equipment list for the AdV process used in Chapter 6 and 6. All parts of the process use stainless steel equipment. Hence, tanks are installed, requiring SIP/CIP steps.

Table A.11: AdV MU equipment summary by step

		2000		1000		200		50		
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
1	Media Hold Tank	1	11.49	1	5.74	1	1.15	1	0.29	L
1	Media Prep Tank	1	11.49	1	5.74	1	1.15	1	0.29	L
1	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
1	Rocking Bioreactor Skid	1	50.00	1	50.00	1	50.00	1	50.00	L
1	Test Tube Rack	1	0.01	1	0.01	1	0.01	1	0.01	L
2	Media Prep Tank	1	63.37	1	31.69	1	6.34	1	1.58	L
2	Media Hold Tank	1	63.37	1	31.69	1	6.34	1	1.58	L
2	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
2	Rocking Bioreactor Skid	1	200.00	1	200.00	1	200.00	1	200.00	L
3	Air Filter	1	1,444.76	1	722.38	1	144.48	1	36.12	L/h
3	Media Hold Tank	1	482.41	1	241.21	1	48.24	1	12.06	L
3	Medium Hold Tank	1	482.41	1	241.21	1	48.24	1	12.06	L
3	Centrifugal Compressor	1	0.76	1	0.38	1	0.08	1	0.02	kW
3	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
3	Mixer	1	9.69	1	4.84	1	0.97	1	0.24	kg/h
3	Seed Bioreactor	1	349.71	1	174.85	1	34.97	1	8.74	L
4	Air Filter	1	11,792.49	1	5,896.26	1	1,179.25	1	294.81	L/h
4	Bioreactor	1	2,502.46	1	1,251.23	1	250.25	1	62.56	L
4	Media Hold Tank	1	1,905.54	1	952.77	1	190.55	1	47.64	L
4	Media Prep Tank	1	1,905.54	1	952.77	1	190.55	1	47.64	L
4	Centrifugal Compressor	1	6.17	1	3.09	1	0.62	1	0.15	kW
4	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
4	Disposable Generic Container Skid	1	58.00	1	29.00	1	6.00	1	2.00	L
4	Mixer	1	79.06	1	39.53	1	7.91	1	1.98	kg/h
4	Roller Bottle Rack	1	2.20	1	2.20	1	2.20	1	2.20	L
5	Lysis Tank	1	2,798.92	1	1,399.46	1	279.89	1	69.97	L
5	Lysis Hold Tank	1	575.16	1	287.58	1	57.52	1	14.38	L
5	Lysis Buffer Prep Tank	1	575.16	1	287.58	1	57.52	1	14.38	L
5	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
6	Disk-Stack Centrifuge	1	627.57	1	313.78	1	62.76	1	15.69	L/h
7	Dead-End Filter	1	3.50	1	1.75	1	0.35	1	0.09	m2
8	Ultrafiltration Tank	1	2,798.46	1	1,402.36	1	283.50	1	71.62	L

Table A.11: AdV MU equipment summary by step (cont'd)

		2000		1000		200		50		
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
8	Buffer Hold Tank	1	2,139.88	1	1,069.94	1	213.99	1	53.50	L
8	Buffer Prep Tank	1	2,139.88	1	1,069.94	1	213.99	1	53.50	L
8	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
8	Diafilter	1	44.46	1	22.28	1	4.50	1	1.14	m2
9	Wash Buffer Prep Tank	1	2,277.41	1	1,138.70	1	227.74	1	56.94	L
9	Wash Buffer Hold Tank	1	2,277.41	1	1,138.70	1	227.74	1	56.94	L
9	Loading Buffer Hold Tank	1	4,554.81	1	2,277.41	1	455.48	1	113.87	L
9	Elution Hold Tank	1	114.70	1	57.35	1	11.47	1	2.87	L
9	Regen Buffer Tank	1	1,890.90	1	945.45	1	189.09	1	47.27	L
9	Regen Hold Tank	1	1,890.90	1	945.45	1	189.09	1	47.27	L
9	Elution Buffer Prep Tank	1	88.38	1	44.19	1	8.84	1	2.21	L
9	Elution Hold Tank	1	88.38	1	44.19	1	8.84	1	2.21	L
9	Loading Buffer Prep Tank	1	4,554.81	1	2,277.41	1	455.48	1	113.87	L
9	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
9	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
9	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
9	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
9	IEX Column	1	7.65	1	3.83	1	0.77	1	0.19	L
10	Formulation Buffer Prep Tank	1	9,708.16	1	4,854.08	1	970.82	1	242.70	L
10	Buffer Hold Tank	1	9,708.16	1	4,854.08	1	970.82	1	242.70	L
10	Dead-End Filter	1	40.00	1	20.00	1	10.00	1	10.00	m2
10	Diafilter	1	3.08	1	1.54	1	0.31	1	0.08	m2
11	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
12	Freeze-Thaw Module	1	57.36	1	28.68	1	5.74	1	1.43	L
13	Formulation Tank	1	1,169.53	1	584.77	1	116.95	1	29.24	L
13	Dead-End Filter	1	220.00	1	110.00	1	30.00	1	10.00	m2
14	Ultrafilter	1	9.33	1	4.66	1	0.93	1	0.23	m2
15	Assembler	5	350.94	3	292.45	1	175.47	1	43.87	entities/min
15	Discrete Freeze-Thaw Module	17	96.88	9	91.50	2	82.35	1	41.17	L
15	Filler	5	350.94	3	292.45	1	175.47	1	43.87	entities/min
15	Labeler	264	398.79	132	398.79	27	389.93	7	376.00	entities/h
15	Packer	11	382.84	6	350.94	2	210.56	1	105.28	entities/min

A.3.2 LV manufacturing

Table A.12 displays equipment list for the LV process used in Chapter 6 and 6. All parts of the process use stainless steel equipment. Hence, tanks are installed, requiring SIP/CIP steps.

Table A.12: LV MU equipment summary by step

		2000		1000		200		50		
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
1	Air Filter	1	745.77	1	372.89	1	74.58	1	18.64	L/h
1	Medium Hold Tank	1	11.48	1	5.74	1	1.15	1	0.29	L
1	Media Prep Tank	1	11.48	1	5.74	1	1.15	1	0.29	L
1	Centrifugal Compressor	1	0.39	1	0.2	1	0.04	1	0.01	kW
1	Dead-End Filter	1	10	1	10	1	10	1	10	m2
1	Mixer	1	5	1	2.5	1	0.5	1	0.13	kg/h
1	Rocking Bioreactor Skid	1	50	1	50	1	50	1	50	L
1	Test Tube Rack	1	0.01	1	0.01	1	0.01	1	0.01	L
2	Air Filter	1	496.32	1	248.16	1	49.63	1	12.41	L/h
2	Media Hold Tank	1	63.33	1	31.66	1	6.33	1	1.58	L
2	Media Prep Tank	1	63.33	1	31.66	1	6.33	1	1.58	L
2	Centrifugal Compressor	1	0.26	1	0.13	1	0.03	1	0.01	kW
2	Dead-End Filter	1	10	1	10	1	10	1	10	m2
2	Mixer	1	3.33	1	1.66	1	0.33	1	0.08	kg/h
2	Rocking Bioreactor Skid	1	200	1	200	1	200	1	200	L
3	Media Prep Tank	1	482.07	1	241.04	1	48.21	1	12.05	L
3	Media Hold Tank	1	482.07	1	241.04	1	48.21	1	12.05	L
3	Dead-End Filter	1	10	1	10	1	10	1	10	m2
3	Seed Bioreactor	1	349.46	1	174.73	1	34.95	1	8.74	L
4	Bioreactor	1	2,802.37	1	1,401.18	1	280.24	1	70.06	L
4	Media Hold Tank	1	1,904.19	1	952.09	1	190.42	1	47.61	L
4	Media Prep Tank	1	1,904.19	1	952.09	1	190.42	1	47.61	L
4	Dead-End Filter	1	10	1	10	1	10	1	10	m2
4	Roller Bottle Rack	1	2.2	1	2.2	1	2.2	1	2.2	L
5	Endonuclease Tank	1	2,112.93	1	1,056.96	1	212.18	1	53.8	L
6	Microfilter	1	2,223.10	1	1,111.55	1	222.31	1	55.58	L
6	Microfilter	1	36.19	1	36.19	1	36.19	1	36.19	m2
7	Wash Buffer Hold Tank	1	92.02	1	46.01	1	9.2	1	2.3	L

Table A.12: LV MU equipment summary by step (cnt'd)

		2000		1000		200		50		
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
7	Wash Buffer Prep Tank	1	92.02	1	46.01	1	9.2	1	2.3	L
7	Regen Hold Tank	1	112.35	1	56.18	1	11.24	1	5.62	L
7	Regen Buffer Tank	1	112.35	1	56.18	1	11.24	1	5.62	L
7	Elution Buffer Tank	1	88.32	1	44.16	1	8.83	1	2.21	L
7	Loading Hold Tank	1	92.02	1	46.01	1	9.2	1	2.3	L
7	Loading Buffer Tank	1	92.02	1	46.01	1	9.2	1	2.3	L
7	Elution Hold Tank	1	88.32	1	44.16	1	8.83	1	2.21	L
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	IEX Column	1	1.18	1	0.59	1	0.12	1	0.03	L
8	Ultrafiltration Tank	1	9.19	1	4.6	1	0.92	1	0.23	L
8	Buffer Hold Tank	1	22.29	1	11.14	1	2.23	1	0.56	L
8	Formulation Tank	1	22.29	1	11.14	1	2.23	1	0.56	L
8	Dead-End Filter	1	10	1	10	1	10	1	10	m2
8	Diafilter	1	0.82	1	0.41	1	0.08	1	0.02	m2
8	Disposable Generic Container Skid	1	2	1	1	1	1	1	1	L
9	Dead-End Filter	1	10	1	10	1	10	1	10	m2
10	Dead-End Filter	1	10	1	10	1	10	1	10	m2
12	Storage Tank	1	12.65	1	6.33	1	1.27	1	0.32	L
12	Freeze-Thaw Module	1	1.53	1	0.77	1	0.15	1	0.04	L
12	Ultrafilter	1	0.25	1	0.12	1	0.03	1	0.01	m2
13	Assembler	1	113.93	1	56.96	1	11.39	1	2.85	entities/h
13	Filler	1	1.9	1	0.95	1	0.19	1	0.05	entities/min
13	Labeler	1	113.94	1	56.97	1	11.39	1	2.85	entities/h
13	Packer	1	4.56	1	2.28	1	0.46	1	0.11	entities/min
15	Discrete Freeze-Thaw Module	1	8.45	1	4.22	1	0.85	1	0.21	L

A.4 SU equipment flowsheets

A.4.1 LV manufacturing

Table A.15 displays equipment list for the LV process used in Chapter 6 and 6. All parts of the process use single-use equipment. Hence, skids are installed to support disposable bags.

A.4.2 AdV manufacturing

Table A.13 displays equipment list for the AdV process used in Chapter 6 and 6. All parts of the process use single-use equipment. Hence, skids are installed to support disposable bags.

Table A.13: AdV SU equipment summary by step

		2000		1000		200		50	
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
1	Media Hold Skid	1	100	1	100	1	100	1	100	L
1	Media Prep Skid	1	100	1	100	1	100	1	100	L
1	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
1	Rocking Bioreactor Skid	1	50.00	1	50.00	1	50.00	1	50.00	L
1	Test Tube Rack	1	0.01	1	0.01	1	0.01	1	0.01	L
2	Media Prep Skid	1	100	1	100	1	100	1	100	L
2	Media Hold Skid	1	100	1	100	1	100	1	100	L
2	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
2	Rocking Bioreactor Skid	1	200.00	1	200.00	1	200.00	1	200.00	L
3	Air Filter	1	1,444.76	1	722.38	1	144.48	1	36.12	L/h
3	Media Hold Skid	1	1000	1	1000	1	1000	1	100	L
3	Medium Hold Skid	1	1000	1	1000	1	1000	1	100	L
3	Centrifugal Compressor	1	0.76	1	0.38	1	0.08	1	0.02	kW
3	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
3	Mixer	1	9.69	1	4.84	1	0.97	1	0.24	kg/h
3	Seed Bioreactor	1	349.71	1	174.85	1	34.97	1	8.74	L
4	Air Filter	1	11,792.49	1	5,896.26	1	1,179.25	1	294.81	L/h
4	Bioreactor	1	2,502.46	1	1,251.23	1	250.25	1	62.56	L
4	Media Hold Skid	2	1000	1	1000	1	1000	1	100	L
4	Media Prep Skid	2	1000	1	1000	1	1000	1	100	L
4	Centrifugal Compressor	1	6.17	1	3.09	1	0.62	1	0.15	kW
4	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
4	Disposable Generic Container Skid	1	58.00	1	29.00	1	6.00	1	2.00	L
4	Mixer	1	79.06	1	39.53	1	7.91	1	1.98	kg/h
4	Roller Bottle Rack	1	2.20	1	2.20	1	2.20	1	2.20	L
5	Lysis Skid	1	2,798.92	1	1,399.46	1	279.89	1	69.97	L
5	Lysis Hold Skid	1	1000	1	1000	1	100	1	100	L
5	Lysis Buffer Prep Skid	1	1000	1	1000	1	100	1	100	L
5	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
6	Disk-Stack Centrifuge	1	627.57	1	313.78	1	62.76	1	15.69	L/h
7	Dead-End Filter	1	3.50	1	1.75	1	0.35	1	0.09	m2
8	Ultrafiltration Skid	1	2,798.46	1	1,402.36	1	283.50	1	71.62	L

Table A.14: AdV SU equipment summary by step (ctn'd)

		2000		1000		200		50	
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
8	Buffer Hold Skid	3	1000	2	1000	1	1000	1	100	L
8	Buffer Prep Skid	3	1000	2	1000	1	1000	1	100	L
8	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
8	Diafilter	1	44.46	1	22.28	1	4.50	1	1.14	m2
9	Wash Buffer Prep Skid	3	1000	2	1000	1	1000	1	100	L
9	Wash Buffer Hold Skid	3	1000	2	1000	1	1000	1	100	L
9	Loading Buffer Hold Skid	3	1000	3	1000	1	1000	1	1000	L
9	Elution Hold Skid	1	1000	1	100	1	100	1	100	L
9	Regen Buffer Skid	2	1000	1	1000	1	1000	1	100	L
9	Regen Hold Skid	2	1000	1	1000	1	1000	1	100	L
9	Elution Buffer Prep Skid	1	88.38	1	44.19	1	8.84	1	2.21	L
9	Elution Hold Skid	1	88.38	1	44.19	1	8.84	1	2.21	L
9	Loading Buffer Prep Skid	1	4,554.81	1	2,277.41	1	455.48	1	113.87	L
9	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
9	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
9	Dead-End Filter	1	10.00	1	10.00	1	10.00	1	10.00	m2
9	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
9	IEX Column	1	7.65	1	3.83	1	0.77	1	0.19	L
10	Formulation Buffer Prep Skid	8	1000	1	1000	1	1000	1	1000	L
10	Buffer Hold Skid	8	1000	4	1000	1	1000	1	1000	L
10	Dead-End Filter	1	40.00	1	20.00	1	10.00	1	10.00	m2
10	Diafilter	1	3.08	1	1.54	1	0.31	1	0.08	m2
11	Dead-End Filter	1	20.00	1	10.00	1	10.00	1	10.00	m2
12	Freeze-Thaw Module	1	57.36	1	28.68	1	5.74	1	1.43	L
13	Formulation Skid	2	1000	1	1000	1	1000	1	29.24	L
13	Dead-End Filter	1	220.00	1	110.00	1	30.00	1	10.00	m2
14	Ultrafilter	1	9.33	1	4.66	1	0.93	1	0.23	m2
15	Assembler	5	350.94	3	292.45	1	175.47	1	43.87	entities/min
15	Discrete Freeze-Thaw Module	17	96.88	9	91.50	2	82.35	1	41.17	L
15	Filler	5	350.94	3	292.45	1	175.47	1	43.87	entities/min
15	Labeler	264	398.79	132	398.79	27	389.93	7	376.00	entities/h
15	Packer	11	382.84	6	350.94	2	210.56	1	105.28	entities/min

Table A.15: LV SU equipment summary by step

		2000		1000		200		50		
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
1	Air Filter	1	745.77	1	372.89	1	74.58	1	18.64	L/h
1	Medium Hold Skid	1	100	1	100	1	100	1	100	L
1	Medium Prep Skid	1	100	1	100	1	100	1	100	L
1	Centrifugal Compressor	1	0.39	1	0.2	1	0.04	1	0.01	kW
1	Dead-End Filter	1	10	1	10	1	10	1	10	m2
1	Mixer	1	5	1	2.5	1	0.5	1	0.13	kg/h
1	Rocking Bioreactor Skid	1	50	1	50	1	50	1	50	L
1	Test Tube Rack	1	0.01	1	0.01	1	0.01	1	0.01	L
2	Air Filter	1	496.32	1	248.16	1	49.63	1	12.41	L/h
2	Media Hold Skid	1	100	1	31.66	1	100	1	100	L
2	Media Prep Skid	1	100	1	31.66	1	100	1	100	L
2	Centrifugal Compressor	1	0.26	1	0.13	1	0.03	1	0.01	kW
2	Dead-End Filter	1	10	1	10	1	10	1	10	m2
2	Mixer	1	3.33	1	1.66	1	0.33	1	0.08	kg/h
2	Rocking Bioreactor Skid	1	200	1	200	1	200	1	200	L
3	Media Prep Skid	1	1000	1	1000	1	100	1	100	L
3	Media Hold Skid	1	1000	1	1000	1	100	1	100	L
3	Dead-End Filter	1	10	1	10	1	10	1	10	m2
3	Seed Bioreactor	1	349.46	1	174.73	1	34.95	1	8.74	L
4	Bioreactor	1	2,802.37	1	1,401.18	1	280.24	1	70.06	L
4	Media Hold Skid	1	1000	1	1000	1	1000	1	100	L
4	Media Prep Skid	1	1000	1	1000	1	1000	1	100	L
4	Dead-End Filter	1	10	1	10	1	10	1	10	m2
4	Roller Bottle Rack	1	2.2	1	2.2	1	2.2	1	2.2	L
5	Endonuclease Skid	1	2,112.93	1	1,056.96	1	212.18	1	53.8	L
6	Microfilter	1	2,223.10	1	1,111.55	1	222.31	1	55.58	L
6	Microfilter	1	36.19	1	36.19	1	36.19	1	36.19	m2
7	Wash Buffer Hold Skid	1	92.02	1	46.01	1	9.2	1	2.3	L

Table A.16: LV SU equipment summary by step (ctn'd)

		2000		1000		200		50		
Step	**Type**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	**#**	**Capacity**	
7	Wash Buffer Prep Skid	1	100	1	100	1	100	1	100	L
7	Regen Hold Skid	1	1000	1	100	1	100	1	100	L
7	Regen Buffer Skid	1	1000	1	100	1	100	1	100	L
7	Elution Buffer Skid	1	100	1	100	1	100	1	100	L
7	Loading Hold Skid	1	100	1	100	1	100	1	100	L
7	Loading Buffer Skid	1	100	1	100	1	100	1	100	L
7	Elution Hold Skid	1	88.32	1	44.16	1	8.83	1	2.21	L
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	Dead-End Filter	1	10	1	10	1	10	1	10	m2
7	IEX Column	1	1.18	1	0.59	1	0.12	1	0.03	L
8	Ultrafiltration Skid	1	9.19	1	4.6	1	0.92	1	0.23	L
8	Buffer Hold Skid	1	100	1	100	1	100	1	100	L
8	Formulation Skid	1	100	1	100	1	100	1	100	L
8	Dead-End Filter	1	10	1	10	1	10	1	10	m2
8	Diafilter	1	0.82	1	0.41	1	0.08	1	0.02	m2
8	Disposable Generic Container Skid	1	2	1	1	1	1	1	1	L
9	Dead-End Filter	1	10	1	10	1	10	1	10	m2
10	Dead-End Filter	1	10	1	10	1	10	1	10	m2
12	Storage Skid	1	12.65	1	6.33	1	1.27	1	0.32	L
12	Freeze-Thaw Module	1	1.53	1	0.77	1	0.15	1	0.04	L
12	Ultrafilter	1	0.25	1	0.12	1	0.03	1	0.01	m2
13	Assembler	1	113.93	1	56.96	1	11.39	1	2.85	entities/h
13	Filler	1	1.9	1	0.95	1	0.19	1	0.05	entities/min
13	Labeler	1	113.94	1	56.97	1	11.39	1	2.85	entities/h
13	Packer	1	4.56	1	2.28	1	0.46	1	0.11	entities/min
15	Discrete Freeze-Thaw Module	1	8.45	1	4.22	1	0.85	1	0.21	L

Appendix B

Global sensitivity analysis

B.1 GSA I & GSA II results

Chapter 3 displays GSA results as plots of the highest 5 ranking input parameters. In this Section, the full GSA plots obtained through the sequential approach are presented. These include the results of GSA I to identify key operational inputs and GSA II to identify combined operational and cost inputs for AdV (Figure B.1), LV (Figure B.2), AAV (Figure B.3). In addition, the full input set considered in GSA III is displayed for the AdV case (Figure B.4), LV (Figure B.5), AAV (Figure B.6).

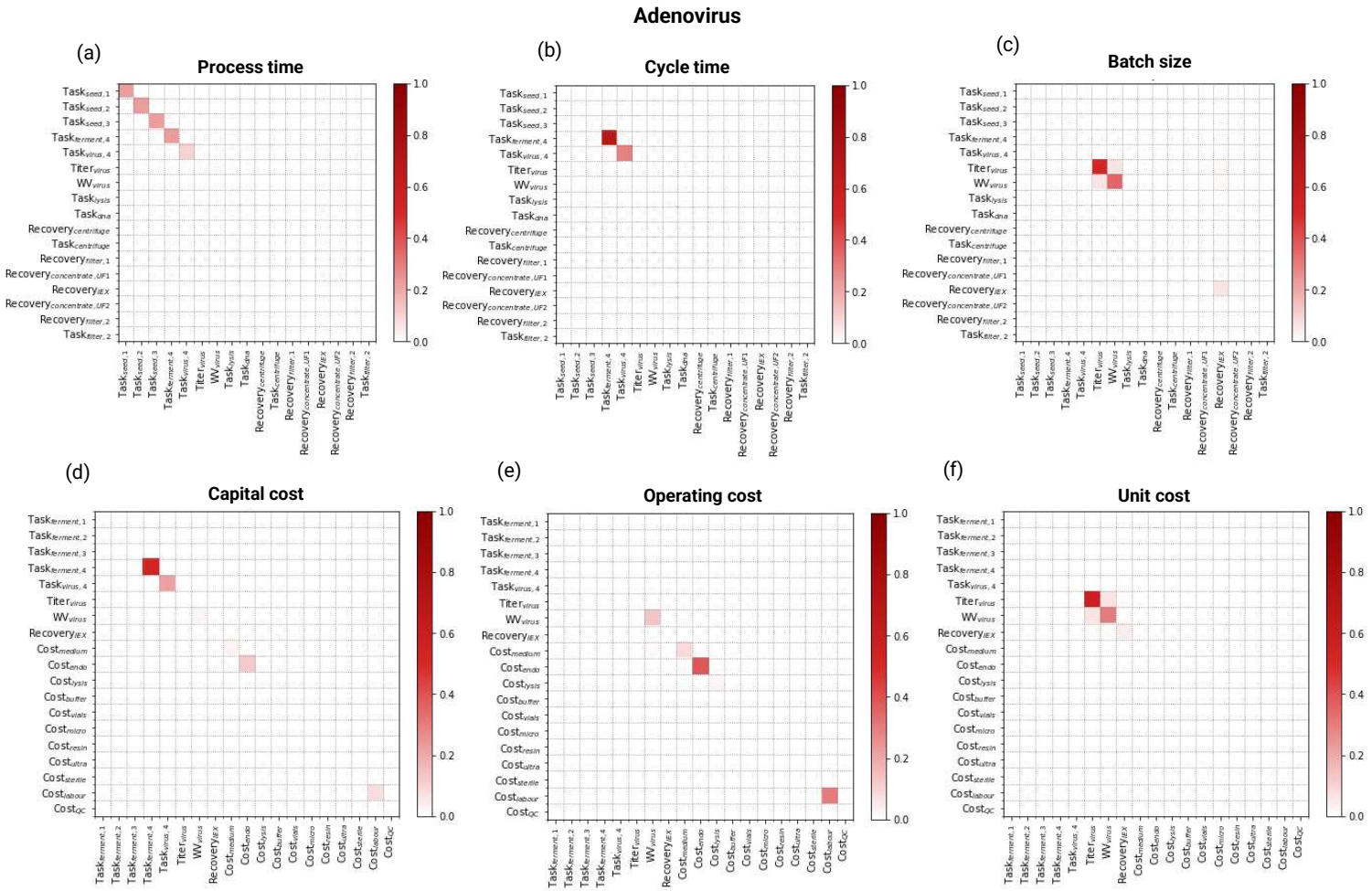


Figure B.1: GSA I-II AdV results

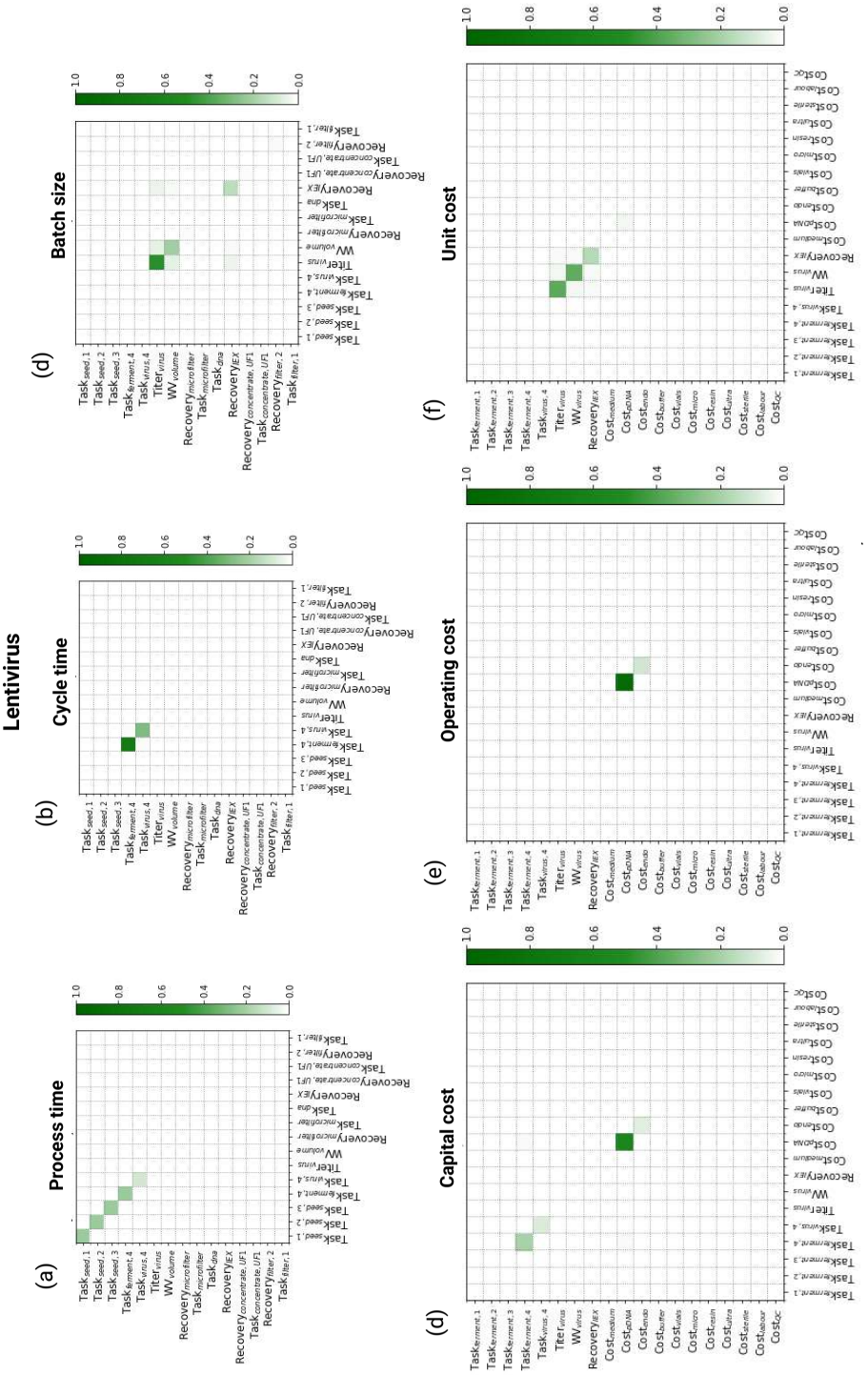


Figure B.2: GSA I-II LV results

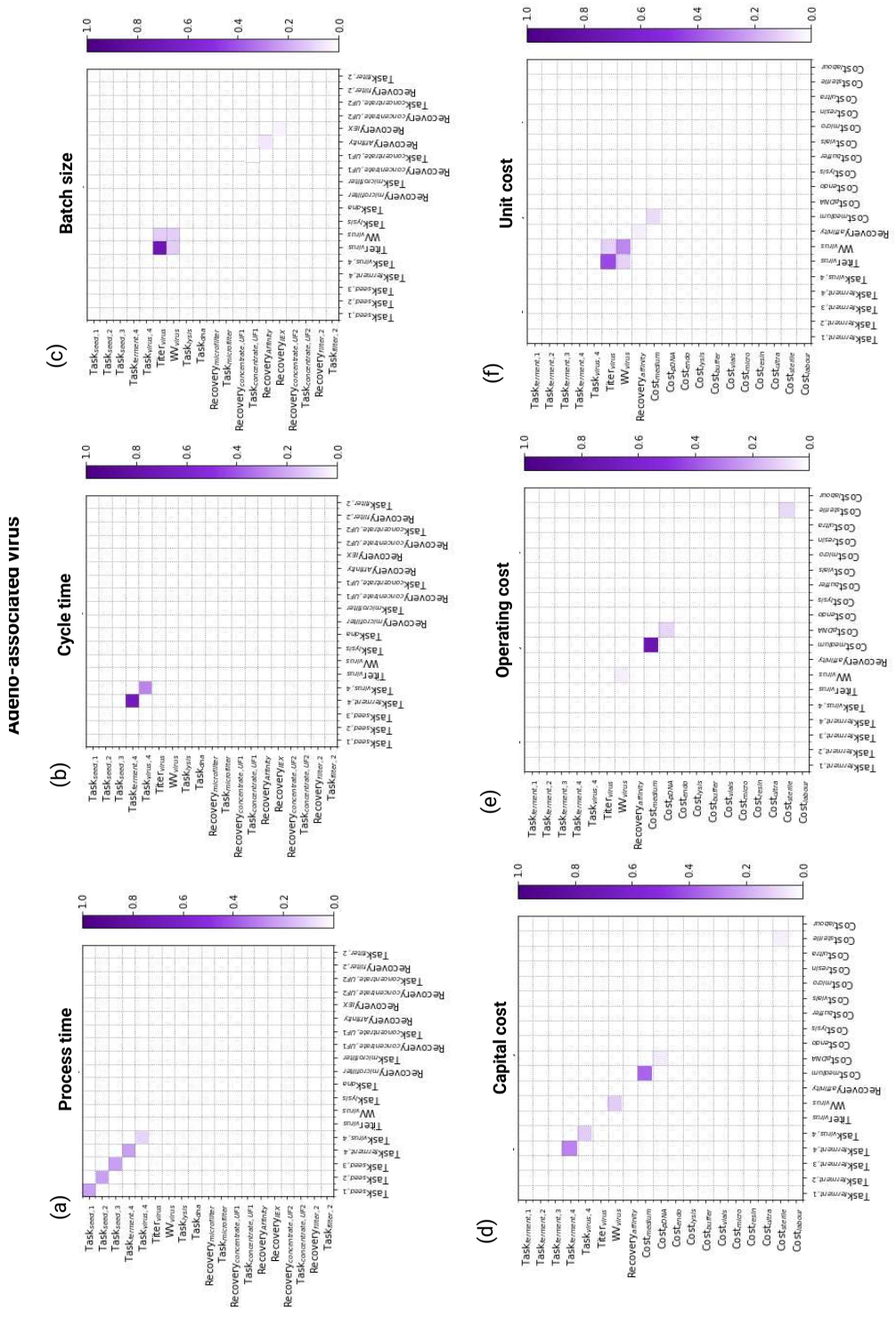


Figure B.3: GSA I-II AAV results

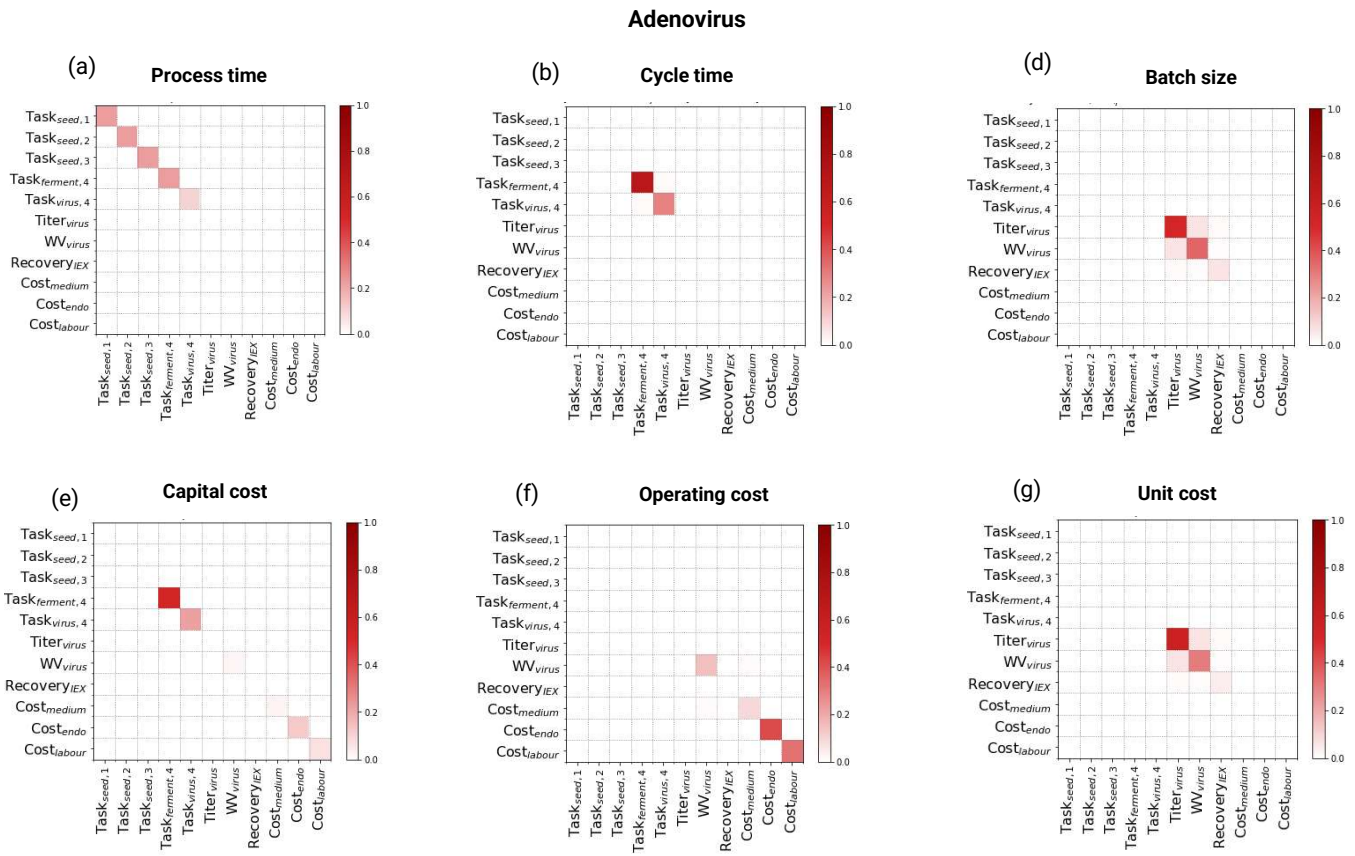


Figure B.4: GSA III AdV results

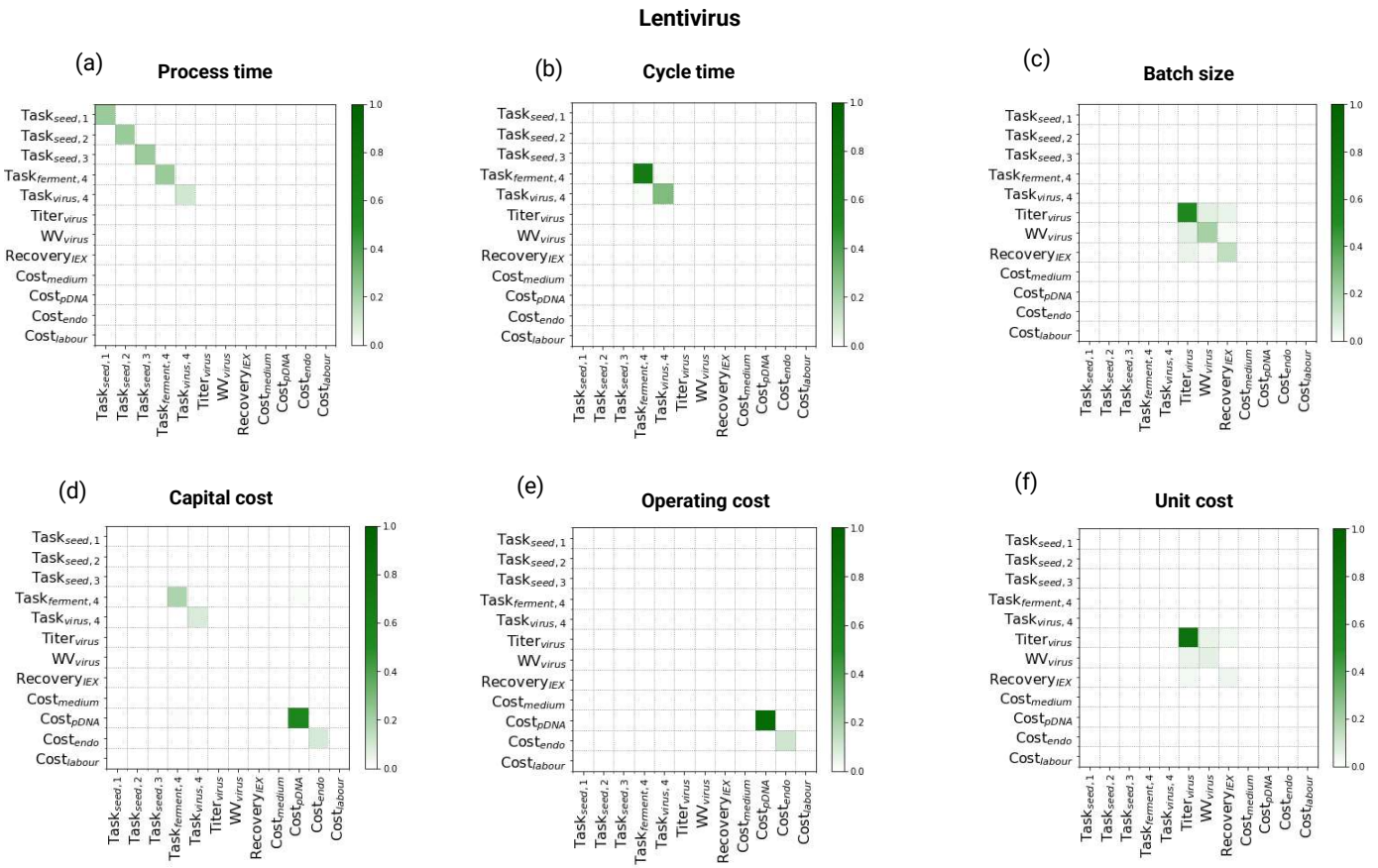


Figure B.5: GSA III LV results

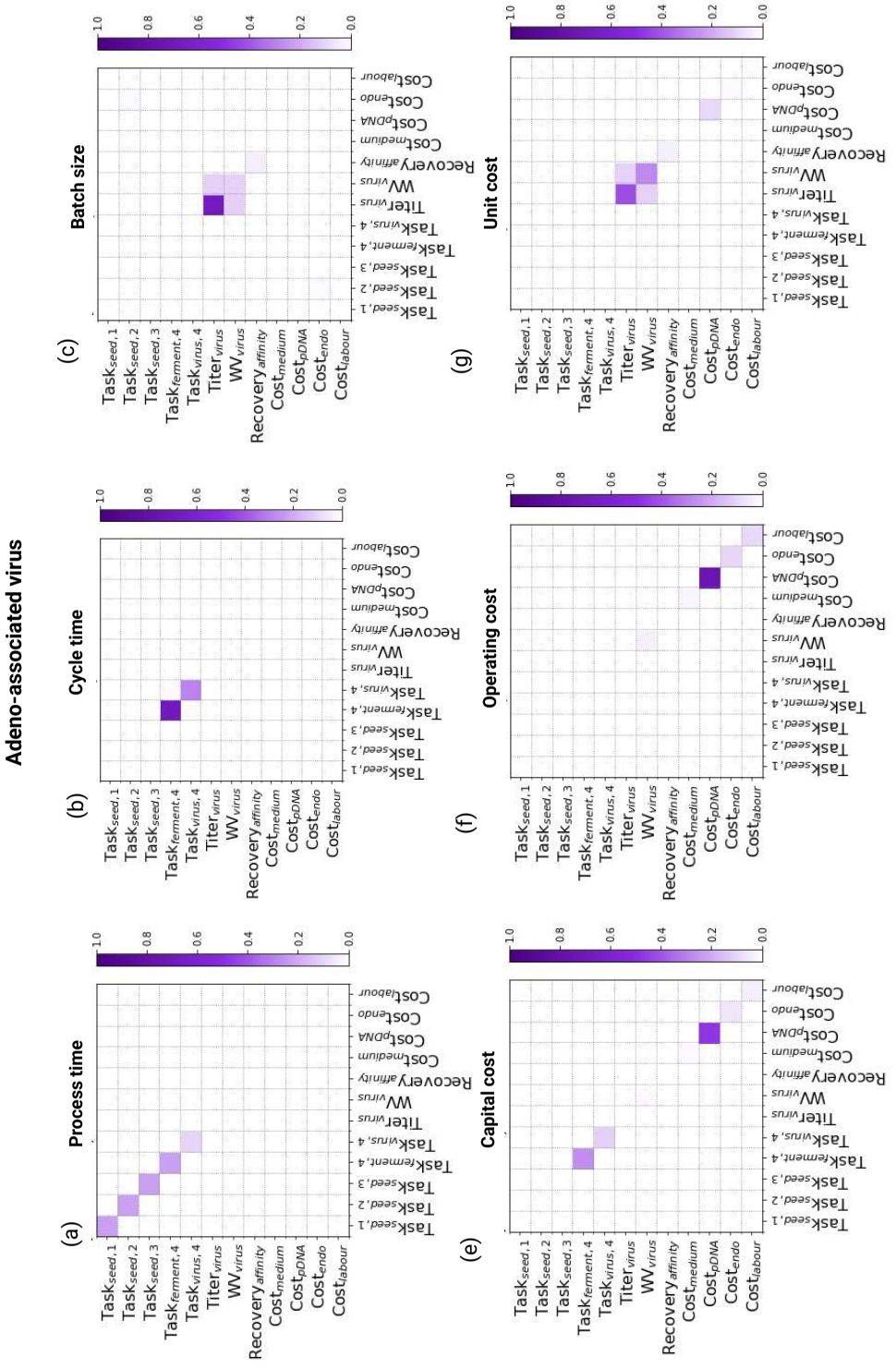


Figure B.6: GSA III AAV results

Appendix C

Optimisation inputs

C.1 Global supply chain parameters

Optimisation input parameters which are independent of the case study and formulation considered are reported below.

Table C.1: Supply chain parameters

Q^{MAX}	5.00E+08	doses
Q^{MIN}	0	doses
T^{MIN}	0	days
T^{MAX}	330	days
U^{MAX}	330	days
c^{OPk}	0.4	$/batch
c^{T}	0.05	$/dose/km

Table C.2: Transport distance ρ between upstream node $n \in N^J$ and storage node $n \in N^K$

Upstream node	Distance (km)
j1	538
j2	799
j3	984

Table C.3: Transport distance ρ between storage node $n \in N^K$ and demand zone upstream node $n \in N^Z$ and

Downstream node	Distance (km)
l1	1640
l2	1878
l3	632
l4	905
l5	1140

C.2 Parameters Chapter 4 and 5

Chapter 4 and 5 consider the integration of the empirical samples representing uncertainty generated through the uncertainty analyses in Chapter 3. The inputs considered in the optimisation are reported below, for the deterministic cases.

C.2.1 Deterministic cases

Table C.4: Capital costs coefficients c^{CAP} ($/line) for every section s and scale a for the deterministic cases *best, median, worst*

Section	Scale	Best	Median	Worst
USP	50	4.03E+07	4.04E+07	4.04E+07
	200	5.58E+07	5.60E+07	5.61E+07
	1000	7.22E+07	7.24E+07	7.28E+07
	2000	8.31E+07	8.40E+07	8.54E+07
DSP	50	2.32E+07	2.32E+07	2.32E+07
	200	2.40E+07	2.40E+07	2.41E+07
	1000	2.95E+07	3.02E+07	3.10E+07
	2000	3.55E+07	3.64E+07	3.85E+07
FF	50	2.32E+07	7.76E+06	7.76E+06
	200	2.40E+07	7.76E+06	7.76E+06
	1000	2.95E+07	7.76E+06	7.76E+06
	2000	3.55E+07	7.76E+06	7.76E+06

Table C.5: Facility dependent costs coefficients c^{FAC} ($/line) for every section s and scale a for the deterministic cases *best, median, worst*

Section	Scale	Best	Median	Worst
USP	50	5.56E+06	5.56E+06	5.56E+06
	200	7.72E+06	7.72E+06	7.72E+06
	1000	9.93E+06	9.93E+06	9.93E+06
	2000	1.14E+07	1.15E+07	1.16E+07
DSP	50	3.22E+06	3.22E+06	3.22E+06
	200	3.33E+06	3.33E+06	3.33E+06
	1000	4.08E+06	4.17E+06	4.26E+06
	2000	4.88E+06	4.97E+06	5.25E+06
FF	50	1.13E+06	1.13E+06	1.13E+06
	200	1.13E+06	1.13E+06	1.13E+06
	1000	1.13E+06	1.13E+06	1.13E+06
	2000	1.13E+06	1.13E+06	1.13E+06

Table C.6: Variable costs coefficients c^{VAR} ($/batch) for every section s and scale a for the deterministic cases *best, median, worst*

Section	Scale	Best	Median	Worst
USP	50	8.06E+03	1.61E+04	3.22E+04
	200	2.06E+04	4.11E+04	8.23E+04
	1000	8.85E+04	1.75E+05	3.44E+05
	2000	1.73E+05	3.42E+05	6.74E+05
DSP	50	5.34E+03	1.07E+04	2.14E+04
	200	9.87E+03	1.98E+04	3.96E+04
	1000	3.47E+04	6.85E+04	1.33E+05
	2000	6.56E+04	1.30E+05	2.57E+05
FF	50	5.06E+02	1.01E+03	2.02E+03
	200	5.23E+02	1.05E+03	2.09E+03
	1000	6.23E+02	1.23E+03	2.42E+03
	2000	7.40E+02	1.46E+03	2.87E+03

Table C.7: Labour costs coefficients c^{LAB} ($/batch) for every section s and scale a for the deterministic cases *best, median, worst*

Section	Scale	Best	Median	Worst
USP	50	3.86E+04	7.86E+04	1.37E+05
	200	3.87E+04	7.85E+04	1.37E+05
	1000	3.98E+04	8.01E+04	1.39E+05
	2000	4.05E+04	8.13E+04	1.42E+05
DSP	50	9.44E+03	1.89E+04	3.87E+04
	200	9.48E+03	1.94E+04	3.97E+04
	1000	1.09E+04	2.19E+04	4.35E+04
	2000	1.26E+04	2.50E+04	4.95E+04
FF	50	6.55E+02	1.31E+03	2.63E+03
	200	6.55E+02	1.31E+03	2.63E+03
	1000	6.67E+02	1.32E+03	2.60E+03
	2000	6.72E+02	1.33E+03	2.61E+03

Table C.8: Batch size χ (doses/batch) for each scale a for the deterministic cases *best, median, worst*

Scale	Best	Median	Worst
50	16	5	1
200	63	21	5
1000	317	107	23
2000	635	215	47

Table C.9: Bottleneck time τ (days/batch) for each section s for the deterministic cases *best, median, worst*

Section	Best	Median	Worst
USP	4.7	6.8	9.2
DSP	1.0	1.0	1.0
FF	1.0	1.0	1.0

C.2.2 Stochastic cases

In the stochastic case each of the above parameters listed in Section C.2.1 gains the dimension p and is assigned a value based on the integer n value generated

Table C.10: Batch time α (days/batch) for each section s for the deterministic cases *best, median, worst*

Section	Best	Median	Worst
USP	15.7	20.4	25.5
DSP	0.2	0.2	0.2
FF	0.2	0.2	0.2

either via random or quasi-random sampling. Given the empirical samples for each process generated via UA, the integer n generated is used to retrieve the value of process parameters ϵ which are jointly distributed.

C.3 Parameters Chapter 6

In Chapter 6, techno-economic process input data is split into uncertain throughput quantities and certain cost-related inputs. For both MU and SU cases the process parameters are considered uncertain. The tables below report the values assumed in the first stage for the median and worst-case. For the second stage onwards, the deterministic value is considered.

Table C.11: Batch size χ (doses/batch) for each scale a for each case

Scale	Median	Worst
50	5	1
200	21	5
1000	107	23
2000	215	47

Table C.12: Bottleneck time τ (days/batch) for each section s for each case

Section	Median	Worst
USP	6.8	9.2
DSP	1.0	1.0
FF	1.0	1.0

Table C.13: Batch time α (days/batch) for each section s for each case

Section	Median	Worst
USP	20.4	25.5
DSP	0.2	0.2
FF	0.2	0.2

C.3.1 MU case

Table C.14: MU facility cost coefficients c^{FACjl} ($/line) for process section s, product i and scale a.

Scale	Section	Product	
		LV	AdV
50	USP	8.41E+06	8.34E+06
	DSP	7.39E+06	9.79E+06
	FF	1.34E+06	1.69E+06
200	USP	1.07E+07	1.04E+07
	DSP	7.44E+06	1.04E+07
	FF	1.34E+06	2,647,993
1000	USP	1.33E+07	1.33E+07
	DSP	7.69E+06	1.26E+07
	FF	1.34E+06	9.12E+06
2000	USP	1.50E+07	1.49E+07
	DSP	7.81E+06	1.43E+07
	FF	1.34E+06	1.69E+07

Table C.15: MU variable cost coefficients c^{LABjl} (\$/batch) for process section s, product i and scale a.

Scale	Section	Product	
		LV	AdV
50	USP	7.32E+04	7.73E+04
	DSP	1.88E+04	1.59E+04
	FF	1.48E+03	1.65E+03
200	USP	7.34E+04	7.75E+04
	DSP	1.90E+04	1.65E+04
	FF	1.48E+03	1.74E+03
1000	USP	7.44E+04	7.86E+04
	DSP	2.01E+04	1.98E+04
	FF	1.49E+03	2.13E+03
2000	USP	7.73E+04	7.99E+04
	DSP	2.44E+04	2.39E+04
	FF	1.49E+03	2.62E+03

Table C.16: MU labour cost coefficients c^{LABjl} (\$/batch) for process section s, product i and scale a.

Scale	Section	Product	
		LV	AdV
50	USP	1.38E+04	5.49E+03
	DSP	1.66E+04	1.40E+04
	FF	1.01E+03	9.96E+03
200	USP	3.88E+04	5.52E+03
	DSP	1.87E+04	2.63E+04
	FF	1.05E+03	3.89E+04
1000	USP	1.72E+05	5.64E+03
	DSP	6.46E+04	9.09E+04
	FF	1.23E+03	1.90E+05
2000	USP	3.40E+05	5.77E+03
	DSP	1.22E+05	1.75E+05
	FF	1.46E+03	3.81E+05

C.3.2 SU case

Table C.17: SU facility cost coefficients c^{FACjl} ($/line) for process section s, product i and scale a.

Scale	Section	Product	
		LV	AdV
50	USP	5.56E+06	6.01E+06
	DSP	3.06E+06	4.05E+06
	FF	1.18E+06	1.53E+06
200	USP	5.80E+06	6.12E+06
	DSP	7.44E+06	4.86E+06
	FF	1.34E+06	2.49E+06
1000	USP	7.00E+06	6.95E+06
	DSP	3.63E+06	9.35E+06
	FF	1.18E+06	9.02E+06
2000	USP	6.71E+06	8.51E+06
	DSP	4.27E+06	7.96E+06
	FF	1.18E+06	1.68E+07

C.4 Parameters Chapter 7

The inputs cost and impact-related inputs considered in the optimisation are presented below.

C.4.1 Costs

In this case, the optimisation considers parameters which are defined for every process section s, equipment type e and scale a.

Table C.18: SU labour cost coefficients c^{LABjl} (\$/batch) for process section s, product i and scale a.

Scale	Section	Product	
		LV	AdV
50	USP	7.14E+04	7.65E+04
	DSP	1.46E+04	1.07E+04
	FF	1.31E+03	1.65E+03
200	USP	7.34E+04	5.20E+04
	DSP	1.90E+04	1.10E+04
	FF	1.48E+03	1.74E+03
1000	USP	6.28E+04	7.74E+04
	DSP	1.33E+04	1.18E+04
	FF	1.15E+03	1.90E+03
2000	USP	7.36E+04	7.74E+04
	DSP	1.53E+04	1.54E+04
	FF	1.33E+03	2.15E+03

Table C.19: SU variable cost coefficients c^{LABjl} (\$/batch) for process section s, product i and scale a.

Scale	Section	Product	
		LV	AdV
50	USP	2.62E+04	2.02E+04
	DSP	1.42E+04	1.89E+04
	FF	1.35E+03	1.03E+04
200	USP	3.88E+04	2.02E+04
	DSP	1.87E+04	3.41E+04
	FF	1.05E+03	3.92E+04
1000	USP	1.99E+05	2.71E+04
	DSP	7.26E+04	1.10E+05
	FF	1.57E+03	1.79E+05
2000	USP	3.58E+05	3.80E+04
	DSP	1.34E+05	1.80E+05
	FF	1.80E+03	3.58E+05

Table C.20: Facility cost coefficients c^{FACjl} ($/line) for process section s, equipment type e and scale a.

Scale	Section	Equipment	
		MU	SU
50	USP	8.41E+06	5.56E+06
	DSP	7.39E+06	3.06E+06
	FF	1.34E+06	1.18E+06
200	USP	1.07E+07	5.80E+06
	DSP	7.44E+06	7.44E+06
	FF	1.34E+06	1.34E+06
1000	USP	1.33E+07	7.00E+06
	DSP	7.69E+06	3.63E+06
	FF	1.34E+06	1.18E+06
2000	USP	1.50E+07	6.71E+06
	DSP	7.81E+06	4.27E+06
	FF	1.34E+06	1.18E+06

Table C.21: Labour cost coefficients c^{LABjl} ($/batch) for process section s, equipment type e and scale a.

Scale	Section	Equipment	
LAB		MU	SU
50	USP	7.32E+04	7.14E+04
	DSP	1.88E+04	1.46E+04
	FF	1.48E+03	1.31E+03
200	USP	7.34E+04	7.34E+04
	DSP	1.90E+04	1.90E+04
	FF	1.48E+03	1.48E+03
1000	USP	7.44E+04	6.28E+04
	DSP	2.01E+04	1.33E+04
	FF	1.49E+03	1.15E+03
2000	USP	7.73E+04	7.36E+04
	DSP	2.44E+04	1.53E+04
	FF	1.49E+03	1.33E+03

Table C.22: Variable cost coefficients c^{VARjl} ($/batch) for process section s, equipment type e and scale a.

Scale	Section	Equipment	
		MU	SU
50	USP	1.38E+04	2.62E+04
	DSP	1.66E+04	1.42E+04
	FF	1.01E+03	1.35E+03
200	USP	3.88E+04	3.88E+04
	DSP	1.87E+04	1.87E+04
	FF	1.05E+03	1.05E+03
1000	USP	1.72E+05	1.99E+05
	DSP	6.46E+04	7.26E+04
	FF	1.23E+03	1.57E+03
2000	USP	3.40E+05	3.58E+05
	DSP	1.22E+05	1.34E+05
	FF	1.46E+03	1.80E+03

C.4.2 Impacts

The normalisation factors used to obtained a monetised environmental footprints are included in Table C.23

Table C.23: Normalisation factors μ_f for each endpoint indicator ep

Endpoint	ReCiPe 2016 units	Equivalent USD2017
HH	DALY	1.46E05
ED	species.yr	3.9E07
RA	USD2013	1.1

The optimisation considers parameters which are defined for every process section s, equipment type e, scale a and impact category ep and mp (Table C.24, C.25, C.26 and C.27.

Table C.24: Impacts per batch e^{MP} and e^{EP} for equipment type e process section s and scale $a = 50$

		SU			MU			
		USP	DSP	FF	USP	DSP	FF	Unit
Midpoint	*FPMF*	2.94E-02	1.69E-02	5.23E-03	5.42E-01	2.89E-01	1.39E-02	kg PM2.5 eq
	FRS	9.57E+00	6.57E+00	3.66E-01	4.48E+01	2.68E+01	9.02E-01	kg oil eq
	FECO	1.16E+00	4.27E-01	1.14E-01	1.77E+01	9.23E+00	2.95E-01	kg 1,4-DCB
	FWEU	8.79E-03	4.74E-03	5.41E-04	2.10E+01	1.04E+01	1.49E-03	kg P eq
	GW	3.05E+01	1.89E+01	1.53E+00	1.38E+02	8.28E+01	3.93E+00	kg CO2 eq
	HCT	1.23E+00	4.71E-01	5.45E-01	7.43E+01	3.68E+01	1.44E+00	kg 1,4-DCB
	HNCT	2.28E+01	1.15E+01	2.13E+00	3.75E+02	2.02E+02	5.77E+00	kg 1,4-DCB
	IR	7.80E+00	8.03E+00	6.33E-02	1.82E+01	1.46E+01	1.69E-01	kBq Co-60 eq
	LU	1.88E+01	3.75E+00	3.22E-01	6.56E+01	2.84E+01	8.54E-01	m2a crop eq
	MECO	1.53E+00	5.99E-01	1.63E-01	2.38E+01	1.24E+01	4.22E-01	kg 1,4-DCB
	MEU	2.83E-03	6.45E-04	5.28E-05	4.73E-02	3.51E-02	6.37E-04	kg N eq
	MRS	1.96E-01	4.56E-02	1.47E-01	7.84E+00	3.76E+00	3.87E-01	kg Cu eq
	OFHH	5.11E-02	3.54E-02	4.07E-03	4.23E-01	2.45E-01	1.07E-02	kg NOx eq
	OFTE	5.24E-02	3.59E-02	4.18E-03	4.30E-01	2.48E-01	1.10E-02	kg NOx eq
	SOD	3.93E-05	2.72E-05	4.61E-07	8.97E-05	6.79E-05	1.34E-06	kg CFC11 eq
	TA	7.20E-02	4.73E-02	6.10E-03	1.52E+00	8.04E-01	1.63E-02	kg SO2 eq
	TE	1.03E+02	2.44E+01	4.27E+01	1.08E+03	5.16E+02	1.12E+02	kg 1,4-DCB
	WC	2.95E-01	1.24E-01	1.01E-02	1.24E+01	6.18E+00	2.21E-02	m3
Endpoint	HH	5.69E-05	3.27E-05	7.03E-06	1.66E-03	3.55E-04	1.46E-05	DALY
	Eco	2.87E-07	1.06E-07	1.01E-08	1.02E-05	1.20E-06	2.93E-08	species.yr
	RA	3.11E+00	2.07E+00	1.23E-01	9.46E+01	2.96E+01	8.85E-01	USD2013

Table C.25: Impacts per batch e^{MP} and e^{EP} for equipment type e process section s and scale $a = 200$

		SU			MU			
		USP	DSP	FF	USP	DSP	FF	Unit
Midpoint	*FPMF*	4.54E-01	1.45E-01	6.72E-03	1.94E+00	2.26E-01	1.57E-02	kg PM2.5 eq
	FRS	1.60E+02	4.95E+01	1.31E+00	2.54E+02	5.32E+01	1.88E+00	kg oil eq
	FECO	1.86E+01	5.63E+00	2.48E-01	6.94E+01	7.05E+00	4.41E-01	kg 1,4-DCB
	FWEU	1.62E-01	3.05E-02	1.02E-03	5.82E+01	2.85E+01	2.15E-03	kg P eq
	GW	5.10E+02	1.37E+02	3.81E+00	8.50E+02	1.61E+02	6.34E+00	kg CO2 eq
	HCT	1.44E+01	4.14E+00	5.92E-01	2.18E+02	7.62E+00	1.51E+00	kg 1,4-DCB
	HNCT	3.69E+02	1.08E+02	3.92E+00	1.42E+03	1.61E+02	8.08E+00	kg 1,4-DCB
	IR	1.54E+02	1.15E+01	1.23E-01	1.89E+02	2.21E+01	2.37E-01	kBq Co-60 eq
	LU	3.69E+02	2.90E+01	4.34E-01	5.05E+02	3.70E+01	9.90E-01	m2a crop eq
	MECO	2.45E+01	7.57E+00	3.49E-01	9.32E+01	9.44E+00	6.24E-01	kg 1,4-DCB
	MEU	4.38E-02	1.17E-02	1.68E-04	1.97E-01	9.34E-02	1.60E-03	kg N eq
	MRS	1.51E+00	5.88E-01	1.51E-01	2.27E+01	8.40E-01	3.93E-01	kg Cu eq
	OFHH	8.39E-01	3.01E-01	7.05E-03	1.92E+00	3.70E-01	1.41E-02	kg NOx eq
	OFTE	8.56E-01	3.09E-01	7.33E-03	1.95E+00	3.76E-01	1.46E-02	kg NOx eq
	SOD	5.01E-04	5.65E-04	8.10E-07	4.69E-04	4.57E-04	1.94E-06	kg CFC11 eq
	TA	1.20E+00	3.93E-01	9.97E-03	5.33E+00	5.56E-01	2.11E-02	kg SO2 eq
	TE	1.31E+03	2.71E+02	4.82E+01	4.24E+03	4.16E+02	1.19E+02	kg 1,4-DCB
	WC	5.00E+00	1.34E+00	2.82E-02	3.71E+01	3.49E+00	4.10E-02	m3
Endpoint	HH	1.66E-04	4.56E-05	7.61E-06	1.53E-03	7.34E-04	1.93E-05	DALY
	Eco	1.01E-06	1.62E-07	1.15E-08	2.76E-05	1.32E-05	2.82E-08	species.yr
	RA	8.97E+00	3.35E+00	1.81E-01	3.20E+01	1.53E+01	3.58E-01	USD2013

Table C.26: Impacts per batch e^{MP} and e^{EP} for equipment type e process section s and scale $a = 1000$

		SU			MU			
		USP	DSP	FF	USP	DSP	FF	Unit
Midpoint	*FPMF*	4.54E-01	1.45E-01	6.72E-03	1.94E+00	2.26E-01	1.57E-02	kg PM2.5 eq
	FRS	1.60E+02	4.95E+01	1.31E+00	2.54E+02	5.32E+01	1.88E+00	kg oil eq
	FECO	1.86E+01	5.63E+00	2.48E-01	6.94E+01	7.05E+00	4.41E-01	kg 1,4-DCB
	FWEU	1.62E-01	3.05E-02	1.02E-03	5.82E+01	2.85E+01	2.15E-03	kg P eq
	GW	5.10E+02	1.37E+02	3.81E+00	8.50E+02	1.61E+02	6.34E+00	kg CO2 eq
	HCT	1.44E+01	4.14E+00	5.92E-01	2.18E+02	7.62E+00	1.51E+00	kg 1,4-DCB
	HNCT	3.69E+02	1.08E+02	3.92E+00	1.42E+03	1.61E+02	8.08E+00	kg 1,4-DCB
	IR	1.54E+02	1.15E+01	1.23E-01	1.89E+02	2.21E+01	2.37E-01	kBq Co-60 eq
	LU	3.69E+02	2.90E+01	4.34E-01	5.05E+02	3.70E+01	9.90E-01	m2a crop eq
	MECO	2.45E+01	7.57E+00	3.49E-01	9.32E+01	9.44E+00	6.24E-01	kg 1,4-DCB
	MEU	4.38E-02	1.17E-02	1.68E-04	1.97E-01	9.34E-02	1.60E-03	kg N eq
	MRS	1.51E+00	5.88E-01	1.51E-01	2.27E+01	8.40E-01	3.93E-01	kg Cu eq
	OFHH	8.39E-01	3.01E-01	7.05E-03	1.92E+00	3.70E-01	1.41E-02	kg NOx eq
	OFTE	8.56E-01	3.09E-01	7.33E-03	1.95E+00	3.76E-01	1.46E-02	kg NOx eq
	SOD	5.01E-04	5.65E-04	8.10E-07	4.69E-04	4.57E-04	1.94E-06	kg CFC11 eq
	TA	1.20E+00	3.93E-01	9.97E-03	5.33E+00	5.56E-01	2.11E-02	kg SO2 eq
	TE	1.31E+03	2.71E+02	4.82E+01	4.24E+03	4.16E+02	1.19E+02	kg 1,4-DCB
	WC	5.00E+00	1.34E+00	2.82E-02	3.71E+01	3.49E+00	4.10E-02	m3
Endpoint	*HH*	9.04E-04	2.60E-04	1.07E-05	3.14E-03	3.62E-04	2.27E-05	DALY
	Eco	5.27E-06	8.11E-07	1.94E-08	4.80E-05	2.01E-05	3.66E-08	species.yr
	RA	5.05E+01	1.82E+01	4.94E-01	8.17E+01	1.73E+01	6.76E-01	USD2013

Table C.27: Impacts per batch e^{MP} and e^{EP} for equipment type e process section s and scale $a = 2000$

		SU			MU			
		USP	DSP	FF	USP	DSP	FF	Unit
Midpoint	*FPMF*	8.09E-01	1.76E-01	8.30E-03	2.68E+00	1.09E+00	1.74E-02	kg PM2.5 eq
	FRS	3.01E+02	7.72E+01	2.31E+00	4.10E+02	1.34E+02	2.89E+00	kg oil eq
	FECO	3.49E+01	8.89E+00	3.89E-01	9.27E+01	3.67E+01	5.87E-01	kg 1,4-DCB
	FWEU	3.10E-01	4.34E-02	1.52E-03	7.21E+01	3.58E+01	2.72E-03	kg P eq
	GW	9.55E+02	2.01E+02	6.22E+00	1.30E+03	3.85E+02	8.79E+00	kg CO2 eq
	HCT	2.55E+01	5.11E+00	6.42E-01	2.76E+02	1.30E+02	1.57E+00	kg 1,4-DCB
	HNCT	6.83E+02	1.56E+02	5.81E+00	1.91E+03	7.64E+02	1.02E+01	kg 1,4-DCB
	IR	3.05E+02	1.24E+01	1.89E-01	3.33E+02	3.50E+01	3.02E-01	kBq Co-60 eq
	LU	7.30E+02	4.72E+01	5.52E-01	8.97E+02	1.31E+02	1.12E+00	m2a crop eq
	MECO	4.58E+01	1.19E+01	5.46E-01	1.24E+02	4.94E+01	8.26E-01	kg 1,4-DCB
	MEU	8.18E-02	1.87E-02	2.90E-04	1.69E-01	1.26E-01	2.04E-03	kg N eq
	MRS	2.29E+00	5.32E-01	1.55E-01	2.87E+01	1.33E+01	3.98E-01	kg Cu eq
	OFHH	1.46E+00	3.38E-01	1.02E-02	2.78E+00	9.88E-01	1.74E-02	kg NOx eq
	OFTE	1.49E+00	3.52E-01	1.07E-02	2.83E+00	1.01E+00	1.80E-02	kg NOx eq
	SOD	8.82E-04	1.04E-03	1.18E-06	7.26E-04	9.14E-04	2.41E-06	kg CFC11 eq
	TA	2.17E+00	5.07E-01	1.41E-02	7.30E+00	3.06E+00	2.55E-02	kg SO2 eq
	TE	2.46E+03	4.14E+02	5.40E+01	6.03E+03	2.09E+03	1.25E+02	kg 1,4-DCB
	WC	9.46E+00	2.14E+00	4.73E-02	4.63E+01	2.21E+01	6.23E-02	m3
Endpoint	*HH*	1.66E-03	3.55E-04	1.46E-05	4.35E-03	1.70E-03	2.68E-05	DALY
	Eco	1.02E-05	1.20E-06	2.93E-08	6.27E-05	2.74E-05	4.68E-08	species.yr
	RA	9.46E+01	2.96E+01	8.85E-01	1.31E+02	4.83E+01	1.07E+00	USD2013

Appendix D

Life cycle assessment modelling

D.1 Stainless steel equipment

The volume of the vessels in stainless steel steps relates to a weight which is found in the literature. To obtain values adapted to the size of vessels modelled in the processes in Chapter 3 and 6 and an interpolation is performed based on Tables D.1 continuing values from industrial reports (Sartorius, b).

Table D.1: Stainless steel vessels weight data

Volume (L)	Weight (kg)
50	520
200	630
1000	1225
2000	1415

Table D.2: Stainless steel equipment weight data

Units	Volume (L)	Weight (kg)
Freeze-Thaw	0.01 - 10	600
Discrete Freeze-Thaw	0.01 - 10	600
Rocking bioreactor	50	42.8
Rocking bioreactor	200	197

D.2 Plastics-based equipment

The single-use bags used in SU steps relate to a specific weight which relates to the thickness and material used as reported in industrial reports (Sartorius, a; St. Gobain) based on Tables D.3 and D.4. This allows mapping of material mass to the consumables used in the processes as shown in Tables D.5.

Table D.3: 50mL to 1L Bags

Material	Thickness (mil)	Thickness (mm)
Nylon (+tie)	1.8	0.045
EVOH (+tie)	1.8	0.045
LLDPE	3.4	0.087

Table D.4: 2L to 3000L Bags

Material	Thickness (mil)	Thickness (mm)
Nylon (+tie)	1.8	0.045
EVOH (+tie)	1.8	0.045
LLDPE	8.4	0.214

Table D.5: Bag weight estimates

Items	**Surface Area (m2)**	**Nylon (kg)**	**EVOH (kg)**	**PE (kg)**	**Waste mix (kg)**
FlexBoy Bag 3.0 L	0.140	0.0093	0.0053	0.0340	0.0485
Dft Large Bag	1.446	0.0963	0.0544	0.3518	0.5025
Flexel Bag for Mixing 50 L	0.911	0.0607	0.0343	0.2216	0.3166
Flexel Bag for Mixing 100 L	1.446	0.0963	0.0544	0.3518	0.5025
Flexel Bag for Mixing 1000 L	6.394	0.4258	0.2407	1.5551	2.2216
Flexel Bag for Storage 50 L	0.868	0.0578	0.0327	0.2111	0.3015
Flexel Bag for Storage 100 L	1.377	0.0917	0.0519	0.3350	0.4786
Flexel Bag for Storage 500 L	4.028	0.2682	0.1516	0.9797	1.3995
Flexel Bag for Storage 1000 L	6.394	0.4258	0.2407	1.5551	2.2216
1 L Plastic Bag	0.064	0.0043	0.0024	0.0156	0.0222
50 L Cell Bag	0.868	0.0578	0.0327	0.2111	0.3015
200 L Cell Bag	2.187	0.1456	0.0823	0.5318	0.7598
CultiBag STR 1000	6.394	0.4258	0.2407	1.5551	2.2216
CultiBag STR 200	2.187	0.1456	0.0823	0.5318	0.7598
CultiBag STR 500	4.028	0.2682	0.1516	0.9797	1.3995
Dft Stirred Seed Bioreactor Bag	1.086	0.0723	0.0409	0.2641	0.3773

D.3 LCA inputs and outputs

The input-output inventory flows used to build the LCA models for each process scale are displayed for the MU case (Tables D.6, D.7, D.8 and D.9) and the SU case (Tables D.10, D.11, D.12 and D.13).

Table D.6: MU 2000L process: key input and output flows

		Total	USP	DSP	FF	
	Flow	Amount				
Inputs	calcium chloride	0.15	0.00	0.15	0.00	kg
	carbon dioxide, liquid	154.38	154.38	0.00	0.00	kg
	compressed air, 600 kPa gauge	420.92	420.92	0.00	0.00	m3
	cooling energy	1021.52	920.90	100.62	0.00	MJ
	electricity, high voltage	1243.15	1207.21	35.71	0.23	kWh
	ethylene vinyl acetate copolymer	0.12	0.11	0.00	0.00	kg
	glass tube, borosilicate	0.00	0.00	0.00	0.00	kg
	glucose	13.60	12.95	0.65	0.00	kg
	glycine	5.98	5.98	0.00	0.00	kg
	heat, from steam, in chemical industry	838.16	398.68	439.48	0.00	MJ
	non-ionic surfactant	0.02	0.00	0.02	0.00	kg
	nylon 6	0.21	0.20	0.01	0.00	kg
	oxygen, liquid	14.26	14.26	0.00	0.00	kg
	phosphoric acid, industrial grade, without water, in 85% solution state	333.21	223.00	109.96	0.00	kg
	polyethylene, linear low density, granulate	0.77	0.74	0.03	0.00	kg
	polypropylene, granulate	16.44	4.50	10.81	1.13	kg
	potassium chloride, industrial grade	0.00	0.00	0.00	0.00	kg
	sodium chloride, powder	7.29	0.00	7.21	0.08	kg
	sodium hydroxide, without water, in 50% solution state	93.48	72.52	20.96	0.00	kg
	sodium phosphate	0.81	0.00	0.80	0.01	kg
	steel, chromium steel 18/8, hot rolled	2.47	1.70	0.00	0.77	kg
	storage, 10'000 l	0.00	0.00	0.00	0.00	Item(s)
	TRIS HCl	9.70	0.00	9.70	0.00	kg
	water, deionised	4676.03	17524.41	3707.10	287.93	kg
	water, ultrapure	25219.32	681.00	7685.00	9.91	kg
Outputs	VV Product	359.00	359.00	359.00	359.00	doses
	waste glass	0.00	0.00	0.00	0.00	kg
	waste plastic, mixture	1.11	1.06	0.04	0.00	kg
	waste polypropylene	16.44	4.50	10.81	1.13	kg
	wastewater, average	26.10	15.35	10.47	0.28	m3
	Carbon dioxide, biogenic	17.17	17.17	0.00	0.00	kg

Table D.7: MU 1000L process: key input and output flows

		Total	USP	DSP	FF	
	Flow	Amount				
Inputs	calcium chloride	0.15	0.00	0.15	0.00	kg
	carbon dioxide, liquid	154.38	154.38	0.00	0.00	kg
	compressed air, 600 kPa gauge	420.92	420.92	0.00	0.00	m3
	cooling energy	1021.52	920.90	100.62	0.00	MJ
	electricity, high voltage	1243.15	1207.21	35.71	0.23	kWh
	ethylene vinyl acetate copolymer	0.12	0.11	0.00	0.00	kg
	glass tube, borosilicate	0.00	0.00	0.00	0.00	kg
	glucose	13.60	12.95	0.65	0.00	kg
	glycine	5.98	5.98	0.00	0.00	kg
	heat, from steam, in chemical industry	838.16	398.68	439.48	0.00	MJ
	non-ionic surfactant	0.02	0.00	0.02	0.00	kg
	nylon 6	0.21	0.20	0.01	0.00	kg
	oxygen, liquid	14.26	14.26	0.00	0.00	kg
	phosphoric acid, industrial grade, without water, in 85% solution state	333.21	223.00	109.96	0.00	kg
	polyethylene, linear low density, granulate	0.77	0.74	0.03	0.00	kg
	polypropylene, granulate	16.44	4.50	10.81	1.13	kg
	potassium chloride, industrial grade	0.00	0.00	0.00	0.00	kg
	sodium chloride, powder	7.29	0.00	7.21	0.08	kg
	sodium hydroxide, without water, in 50% solution state	93.48	72.52	20.96	0.00	kg
	sodium phosphate	0.81	0.00	0.80	0.01	kg
	steel, chromium steel 18/8, hot rolled	2.47	1.70	0.00	0.77	kg
	storage, 10'000 l	0.00	0.00	0.00	0.00	Item(s)
	TRIS HCl	9.70	0.00	9.70	0.00	kg
	water, deionised	4676.03	17524.41	3707.10	287.93	kg
	water, ultrapure	25219.32	681.00	7685.00	9.91	kg
Outputs	VV Product	359.00	359.00	359.00	359.00	doses
	waste glass	0.00	0.00	0.00	0.00	kg
	waste plastic, mixture	1.11	1.06	0.04	0.00	kg
	waste polypropylene	16.44	4.50	10.81	1.13	kg
	wastewater, average	26.10	15.35	10.47	0.28	m3
	Carbon dioxide, biogenic	17.17	17.17	0.00	0.00	kg

Table D.8: MU 200L process: key input and output flows

		Total	USP	DSP	FF	
	Flow	Amount				
Inputs	calcium chloride	0.015398	0	0.015398	0	kg
	carbon dioxide, liquid	15.44	15.44	0	0	kg
	compressed air, 600 kPa gauge	42.09219	42.09219	0	0	m3
	cooling energy	115.44	96.04386	19.39615	0	MJ
	electricity, high voltage	157.6154	121.3077	36.28846	0.019231	kWh
	ethylene vinyl acetate copolymer	0.11738	0.114974	0.002407	0	kg
	glass tube, borosilicate	0	1.34	0	0	kg
	glucose	1.36008	1.295155	0.064925	0	kg
	glycine	0.597779	0.597779	0	0	kg
	heat, from steam, in chemical industry	86.14831	40.99769	45.15062	0	MJ
	non-ionic surfactant	0.001601	0	0.001601	0	kg
	nylon 6	0.207673	0.203415	0.004258	0	kg
	oxygen, liquid	1.42	1.42	0	0	kg
	phosphoric acid, industrial grade, without water, in 85% solution state	166.6035	111.625	54.9785	0	kg
	polyethylene, linear low density, granulate	0.758458	0.742906	0.015551	0	kg
	polypropylene, granulate	1.644188	0.45	1.081125	0.113063	kg
	potassium chloride, industrial grade	2.00E-06	0	0	2.00E-06	kg
	sodium chloride, powder	0.729339	0	0.721339	0.008	kg
	sodium hydroxide, without water, in 50% solution state	16.97909	7.252	9.727088	0	kg
	sodium phosphate	0.080972	0	0.07987	0.001102	kg
	steel, chromium steel 18/8, hot rolled	1.712882	0.943651	0	0.769231	kg
	storage, 10'000 l	6.48E-05	3.16E-05	3.31E-05	8.14E-08	Item(s)
	TRIS HCl	0.970393	0	0.970393	0	kg
	water, deionised	2338.015	340.5	3662.821	143.965	kg
	water, ultrapure	10197.35	6533.538	1853.55	0.990896	kg
Outputs	VV Product	36	36	36	36	doses
	waste glass	0	0	0	0	kg
	waste plastic, mixture	1.083511	1.061295	0.022216	0	kg
	waste polypropylene	1.644188	0.45	1.081125	0.113063	kg
	wastewater, average	11	6.16	5.01825	0.133404	m3
	Carbon dioxide, biogenic	1.788462	1.788462	0	0	kg

Table D.9: MU 50L process: key input and output flows

		Total	USP	DSP	FF	
	Flow	Amount				
Inputs	calcium chloride	0.003852	0	0.003852	0	kg
	carbon dioxide, liquid	3.91	3.91	0	0	kg
	compressed air, 600 kPa gauge	10.52269	10.52269	0	0	m3
	cooling energy	35.93152	23.80892	12.1226	0	MJ
	electricity, high voltage	66.54327	30.33173	36.21154	0	kWh
	ethylene vinyl acetate copolymer	0.11738	0.114974	0.002407	0	kg
	glass tube, borosilicate	0	0.335	0.41875	0	kg
	glucose	0.340029	0.323789	0.016241	0	kg
	glycine	0.149445	0.149445	0	0	kg
	heat, from steam, in chemical industry	21.62844	10.26604	11.36241	0	MJ
	non-ionic surfactant	4.00E-04	0	4.00E-04	0	kg
	nylon 6	0.207673	0.203415	0.004258	0	kg
	oxygen, liquid	0.36	0.36	0	0	kg
	phosphoric acid, industrial grade, without water, in 85% solution state	96.63003	64.7425	31.88753	0	kg
	polyethylene, linear low density, granulate	0.758458	0.742906	0.015551	0	kg
	polypropylene, granulate	0.411047	0.1125	0.270281	0.028266	kg
	potassium chloride, industrial grade	5.00E-07	0	0	5.00E-07	kg
	sodium chloride, powder	0.255394	0	0.253394	0.002	kg
	sodium hydroxide, without water, in 50% solution state	7.94192	1.4504	6.49152	0.00E+00	kg
	sodium phosphate	2.03E-02	0	2.00E-02	2.76E-04	kg
	steel, chromium steel 18/8, hot rolled	1.450815	0.681584	0	0.769231	kg
	storage, 10'000 l	1.66E-05	7.89E-06	8.69E-06	2.05E-08	Item(s)
	TRIS HCl	0.24274	0	0.24274	0	kg
	water, deionised	1356.049	197.49	1075.059	84.02	kg
	water, ultrapure	5745.4	3587.451	2157.702	0.247724	kg
Outputs	VV Product	9	9	9	9	doses
	waste glass	0	0	0	0	kg
	waste plastic, mixture	1.083511	1.061295	0.022216	0	kg
	waste polypropylene	0.411047	0.1125	0.270281	0.028266	kg
	wastewater, average	6	2.908558	3	0.084019	m3
	Carbon dioxide, biogenic	0.442308	0.442308	0	0	kg

Table D.10: SU 2000L process: key input and output flows

		Total	USP	DSP	FF	
Input	Flow	Amount				
	calcium chloride	0.153961	0	0.153961	0	kg
	carbon dioxide, liquid	154.38	154.38	0	0	kg
	compressed air, 600 kPa gauge	420.9191	420.9191	0	0	m3
	cooling energy	1061.876	957.218	104.6584	0	MJ
	electricity, high voltage	1243.154	1207.212	35.71154	0.25	kWh
	ethylene vinyl acetate copolymer	2.521709	2.521709	1.557736	0	kg
	glass tube, borosilicate	0	0	0	0	kg
	glucose	13.60092	12.95175	0.649176	0	kg
	glycine	5.977882	5.977882	0	0	kg
	heat, from steam, in chemical industry	3.887139	3.887139	0	0	MJ
	non-ionic surfactant	0.016006	0	0.016006	0	kg
	nylon 6	4.461485	4.461485	2.755994	0	kg
	oxygen, liquid	14.26	14.26	0	0	kg
	phosphoric acid, industrial grade	0	0	0	0	kg
	polyethylene, linear low density, granulate	16.29412	16.29412	10.06537	0	kg
	polypropylene, granulate	4.5	4.5	10.81125	1.130625	kg
	potassium chloride, industrial grade	2E-05	0	0	2.00E-05	kg
	sodium chloride, powder	7.294464	0	7.214464	0.08	kg
	sodium hydroxide	0	0	0	0	kg
	sodium phosphate	0.809916	0	0.798896	0.01102	kg
	steel, chromium steel 18/8, hot rolled	0.937085	0.644778	0	0.292308	kg
	storage, 10'000 l	0.000246	1.20E-04	1.26E-04	3.08E-07	Item(s)
	TRIS HCl	9.702807	0	9.702807	0	kg
	water, deionised	0	0	0	0	kg
	water, ultrapure	2335	1944.177	380.9147	9.90896	kg
Outputs	VV Product	359	359	359	359	doses
	waste glass	0	0	0	0	kg
	waste plastic, mixture	23.27731	23.27731	14.3791	0	kg
	waste polypropylene	4.5	4.5	10.81125	1.130625	kg
	wastewater, average	0	0	0	0	m3
	Carbon dioxide, biogenic	17.84615	17.84615	0	0	kg

Table D.11: SU 1000L process: key input and output flows

		Total	USP	DSP	FF
Input	Flow	Amount			
	calcium chloride	0.076981	0	0.076981	0
	carbon dioxide, liquid	77.18	77.18	0	0
	compressed air, 600 kPa gauge	210.4595	210.4595	0	0
	cooling energy	535.9873	479.0111	56.9762	0
	electricity, high voltage	641.1346	605.0577	36.65385	0.115385
	ethylene vinyl acetate copolymer	2.524115	1.559015	0.965101	0
	glass tube, borosilicate	0	6.7	8.375	0
	glucose	6.800432	6.475841	0.324591	0
	glycine	2.988926	2.988926	0	0
	heat, from steam, in chemical industry	1.926958	1.926958	0	0
	non-ionic surfactant	0.008003	0	0.008003	0
	nylon 6	4.465743	2.758257	1.707486	0
	oxygen, liquid	7.14	7.14	0	0
	phosphoric acid, industrial grade	0	0	0	0
	polyethylene, linear low density, granulate	16.30967	10.07363	6.236036	0
	polypropylene, granulate	8.220938	2.25	5.405625	0.565313
	potassium chloride, industrial grade	1.00E-05	0	0	1.00E-05
	sodium chloride, powder	3.647254	0	3.607254	0.04
	sodium hydroxide	0	0	0	0
	sodium phosphate	0.404958	0	0.399448	0.00551
	steel, chromium steel 18/8, hot rolled	0.830043	0.537736	0	0.292308
	storage, 10'000 l	1.23E-04	5.99E-05	6.28E-05	1.54E-07
	TRIS HCl	4.85141	0	4.85141	0
	water, deionised	0	0	0	0
	water, ultrapure	1172.47	972.0835	195.4323	4.95448
Outputs	VV Product	180	180	180	180
	waste glass	0	0	0	0
	waste plastic, mixture	23.29953	14.3909	8.908623	0
	waste polypropylene	8.220938	2.25	5.405625	0.565313
	wastewater, average	0	0	0	0
	Carbon dioxide, biogenic	8.923077	8.923077	0	0

Table D.12: SU 200L process: key input and output flows

		Total	USP	DSP	FF	
Input	Flow	Amount				
	calcium chloride	0.015398	0	0.015398	0	kg
	carbon dioxide, liquid	15.44	15.44	0	0	kg
	compressed air, 600 kPa gauge	42.09219	42.09219	0	0	m3
	cooling energy	115.44	96.04386	19.39615	0	MJ
	electricity, high voltage	157.6154	121.3077	36.28846	0.019231	kWh
	ethylene vinyl acetate copolymer	0.158259	0.155852	0.002407	0	kg
	glass tube, borosilicate	0	0	0	0	kg
	glucose	1.36008	1.295155	0.064925	0	kg
	glycine	0.597779	0.597779	0	0	kg
	heat, from steam, in chemical industry	0.365458	0.365458	0	0	MJ
	non-ionic surfactant	0.001601	0	0.001601	0	kg
	nylon 6	0.279996	0.275738	0.004258	0	kg
	oxygen, liquid	1.42	1.42	0	0	kg
	phosphoric acid, industrial grade	0	0	0	0	kg
	polyethylene, linear low density, granulate	1.022595	1.007044	0.015551	0	kg
	polypropylene, granulate	1.644188	0.45	1.081125	0.113063	kg
	potassium chloride, industrial grade	2.00E-06	0	0	2.00E-06	kg
	sodium chloride, powder	0.729339	0	0.721339	0.008	kg
	sodium hydroxide	0	0	0	0	kg
	sodium phosphate	0.080972	0	0.07987	0.001102	kg
	steel, chromium steel 18/8, hot rolled	0.650895	0.358587	0	0.292308	kg
	storage, 10'000 l	2.46E-05	1.20E-05	1.26E-05	3.09E-08	Item(s)
	TRIS HCl	0.970393	0	0.970393	0	kg
	water, deionised	0	0	0	0	kg
	water, ultrapure	242.4521	194.4147	47.04647	0.990896	kg
Outputs	VV Product	36	36	36	36	doses
	waste glass	0	0	0	0	kg
	waste plastic, mixture	1.46085	1.438634	0.022216	0	kg
	waste polypropylene	1.644188	0.45	1.081125	0.113063	kg
	wastewater, average	0	0	0	0	m3
	Carbon dioxide, biogenic	1.788462	1.788462	0	0	kg

Table D.13: SU 50L process: key input and output flows

		Total	USP	DSP	FF	
Input	Flow	Amount				
	calcium chloride	0.003852	0	0.003852	0	kg
	carbon dioxide, liquid	3.91	3.91	0	0	kg
	compressed air, 600 kPa gauge	10.52269	10.52269	0	0	m3
	cooling energy	35.93152	23.80892	12.1226	0	MJ
	electricity, high voltage	66.54327	30.33173	36.21154	0	kWh
	ethylene vinyl acetate copolymer	0.158259	0.155852	0.002407	0	kg
	glass tube, borosilicate	0	0.335	0.41875	0	kg
	glucose	0.340029	0.323789	0.016241	0	kg
	glycine	0.149445	0.149445	0	0	kg
	heat, from steam, in chemical industry	0.09967	0.09967	0	0	MJ
	non-ionic surfactant	4.00E-04	0	4.00E-04	0	kg
	nylon 6	0.279996	0.275738	0.004258	0	kg
	oxygen, liquid	0.36	0.36	0	0	kg
	phosphoric acid, industrial grade	0	0	0	0	kg
	polyethylene, linear low density, granulate	1.022595	1.007044	0.015551	0	kg
	polypropylene, granulate	0.411047	0.1125	0.270281	0.028266	kg
	potassium chloride, industrial grade	5.00E-07	0	0	5.00E-07	kg
	sodium chloride, powder	0.255394	0	0.253394	0.002	kg
	sodium hydroxide	0	0	0	0	kg
	sodium phosphate	0.020268	0	0.019992	2.76E-04	kg
	steel, chromium steel 18/8, hot rolled	0.55131	0.259002	0	0.292308	kg
	storage, 10'000 l	6.31E-06	3.00E-06	3.30E-06	7.79E-09	Item(s)
	TRIS HCl	0.24274	0	0.24274	0	kg
	water, deionised	0	0	0	0	kg
	water, ultrapure	70.55479	48.60368	21.70338	0.247724	kg
Outputs	VV Product	9	9	9	9	doses
	waste glass	0	0	0	0	kg
	waste plastic, mixture	1.46085	1.438634	0.022216	0	kg
	waste polypropylene	0.411047	0.1125	0.270281	0.028266	kg
	wastewater, average	0	0	0	0	m3
	Carbon dioxide, biogenic	0.442308	0.442308	0	0	kg