

https://researchcommons.waikato.ac.nz/

Research Commons at the University of Waikato

Copyright Statement:

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

The thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of the thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from the thesis.

Energy System Modelling for the Decarbonisation of Industrial Process Heat

A thesis

submitted in fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Engineering

at

The University of Waikato

by

Daniel Jia Sheng Chong



THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2026

Abstract

Industrial process heat is a major source of fossil fuel use and carbon emissions, yet it is often represented only coarsely in large-scale energy system models. This limits the ability of existing models to assess site-level fuel-switching opportunities, spatial resource allocation, and the wider electricity system implications of industrial decarbonisation. In particular, modelling industrial process heat requires approaches that can capture the diversity of industrial demand sites, the location of fuel and electricity supply options, and the infrastructure needed to connect them. This thesis develops a spatially explicit energy modelling framework to evaluate nationwide fuel-switching pathways for site-level industrial process heat up to approximately 500°C, corresponding to temperatures that can be supplied by high-pressure steam.

The framework combines P-graph superstructure generation with an automated translation tool that converts the structural representation into a Pyomo-based generalised disjunctive programming model. This enables large-scale structural optimisation using reusable modular components and multiple simulation modes. A modular architecture is developed to represent model components in a scalable and reusable manner. The framework is further supported by geospatial preprocessing, standardised input-processing tools, and output reporting modules, which help facilitate the integration of high-spatial-resolution data within a unified P-graph and algebraic modelling workflow.

The methodology is demonstrated through a New Zealand case study that compiles data with high spatial granularity on industrial process heat demand, electricity generation assets, biomass resources, grid infrastructure, and geospatial linkages between supply and demand nodes. The energy systems model is evaluated using three simulation modes: deterministic multiperiod optimisation, sequential deterministic optimisation, and Monte Carlo uncertainty analysis. Together, these modes allow the framework to evaluate systems-level least-cost decarbonisation pathways, grid reinforcement upgrade requirements, and the robustness of system configurations under uncertain input conditions. The deterministic multiperiod factory model captures temporal variation in demand and supply conditions. Building on the multiperiod model, the sequential framework represents staged investment and includes grid reinforcement assessment, while the Monte Carlo analysis evaluates uncertainty in key modelling parameters.

The results indicate that biomass resources play a dominant role in the system-level least-cost industrial process heat decarbonisation in the New Zealand case study. Electrification increases industrial electricity demand by approximately 3.5–9.2 TWh, depending on scenario assumptions, creating additional pressure on generation and grid infrastructure. Sequential analysis shows that grid reinforcement costs can be represented through an adjusted electricity price coefficient, while Monte Carlo analysis provides a distributional view of electrification uptake, biomass utilisation, and spatial outcomes under uncertainty. These findings show that both biomass and electrification should be viewed as complementary decarbonisation pathways in the industrial process heat space, with their relative roles shaped by local resource availability, electricity costs, and infrastructure constraints.

The central contribution of the thesis is a spatially explicit energy modelling framework that combines P-graph superstructure generation with a translation tool converting the structural representation into a Pyomo-based generalised disjunctive programming model. At the same time, the framework demonstrates how P-graph can be extended to support large-scale, spatially explicit, and uncertainty-aware energy system analysis.

AI has been used for proofreading and spelling checks.

Acknowledgments

There were many moments when I doubted my work. Galatians 6:9, which states, “Let us not become weary in doing good, for at the proper time we will reap a harvest if we do not give up,” encouraged me to press on and strive to present meaningful research contributions.

I am deeply grateful to everyone who has supported me through this journey. The thanks that follow are not in order of importance.

Without my family's constant support, my thesis would never have been completed. Mum and Dad, thank you for your sacrifices and always being there for me, even though we were a continent apart. Mum, thank you for giving me home tuition as a child. Dad, thank you for setting yourself as a role model. I have aspired to be a chemical engineer like you since I was young.

I would like to express my sincere gratitude to my chief supervisor, Professor Michael Walmsley, and my second supervisor, Dr Martin Atkins, for their invaluable guidance and continuous support throughout my doctoral studies. Beyond research, working with them also gave me opportunities to experience teaching, lecturing, and guiding students, which has been a meaningful part of my journey.

Although not formally on my supervisory panel, I would also like to express my deep gratitude to my de facto supervisor Dr Timothy Walmsley. He gave me new insights and was always willing to share ideas and engage in discussions that helped redirect my initial research path from what seemed uncertain, towards an area that I genuinely enjoyed exploring.

I am extremely fortunate to have been part of, and funded by, the Ahuora Project, which is comprised of brilliant colleagues who have now become my peers. Special thanks to the PhD students: Ben, Keegan, Lana, Safiya, Matthew, Elsa, Stephen, Jack, who were present when I first started and welcomed me when I arrived in New Zealand for my studies. The supportive environment and shared purpose have made this journey significantly more enriching. Special thanks to Isobella who helped with proofreading my thesis.

I am grateful to Dr Botond Bertók and his colleagues at Széchenyi István University of Győr for hosting me as a visiting researcher for three months and teaching me mathematical modelling techniques for process systems engineering. I must honestly say that without their guidance and willingness to let me pick their brains, I would not have been able to complete the thesis.

Moreover, I would like to thank my friends from New Zealand and Malaysia for their emotional support and for asking about my work even when you don’t understand it, and for keeping me sane throughout my entire doctoral studies.

Nomenclature

Common Abbreviations

Abbreviation	Description
ASCBR	Acceptable Shifting Convergence Band Rule
BEV	Battery Electric Vehicle
BFB	Bubbling Fluidised Bed Reactor
CB	Coal Retrofitted Boiler
COP	Coefficient of Performance
EB	Electrode Boiler
EECA	Energy Efficiency and Conservation Authority
EEUD	Energy End-Use Database
FPVV	Fixed-Price Variable-Volume
GDP	Generalised Disjunctive Programming
GIS	Geographic Information System
GXP	Grid Exit Point
HTHP	High-Temperature Heat Pump
HVDC	High-Voltage Direct Current
ICP	Installation Control Point
IEA	International Energy Agency
LHV	Lower heating value
LP	Linear Programming
LTHP	Low-temperature heat pump
MARKAL	MARKet ALlocation
MBIE	Ministry of Business, Innovation and Employment
MILP	Mixed Integer Linear Programming
MINLP	Mixed-Integer Non-linear Programming
MTHP	Medium-Temperature Heat Pump
NCV	Net Calorific Value
NLP	Non-linear Programming
PNS	Process Network Synthesis
POC	Point of Connection
TAC	Total Annualised Cost
TIMES	The Integrated MARKAL-EFOM System

Parameters

Symbol	Description
$\alpha^{ht,sized}$	Size coefficient for heating technology
cf	Capacity factor
C	Cost coefficient for a component
$\overline{Cap}$	Maximum capacity sizing for a technology component
ε	Half-width of convergence band in ASCBR
f^*	Minimum number of consecutive iterations satisfying convergence criterion
j_{max}	Simulation limit
M	Mapping matrix between layers or technology pathways
η_p	Efficiency of heating technology
$n^{line,max}$	Maximum number of allowable line upgrades
$\overline{P^{tr}}$	Rated capacity of transformer technology
Δt_t	Duration of the time period t

Variables

Symbol	Units	Description
$\alpha^{e\ adjusted}$	-	Cost coefficient adjusted for sales electricity price
B	GJ	Biomass flow
Cap	MW	Capacity of a technology component
CC	NZD	Annualised cost of a component/set
Ch	GWh	Charging for storage
D	GWh	Energy demand
Dis	GWh	Discharging from storage
F	GWh	Energy flow
G^{MeOH}	ton	Methanol production
H	GWh	Heat flow
$n_{i,u,r}^{line}$	-	Number of integer components
Rev^{MeOH}	NZD	Gross sales revenue from green methanol
y	-	Binary variable

Sets

Symbol	Description
T	Set of time periods and index, $t \in T$
J	Set of electricity generation sites and index, $j \in J$
E	Set of electricity generation types and index, $e \in E$
K	Set of biomass resource sites and index, $k \in K$
W	Set of biomass resource types and index, $w \in W$
I	Set of POC sites and index, $i, u \in I$
L	Set of industrial sites and index, $l \in L$
Q	Set of temperature bands and index, $q, q' \in Q$
A^{POC}	Set of POC-to-POC connections and pairs, $(i, u) \in A^{POC}$
$A^{gen,POC}$	Set of generation sites-to-POC connections and pairs, $(j, i) \in A^{gen,POC}$
$A^{POC,fac}$	Set of POC-factory connections and pairs, $(i, l) \in A^{POC,fac}$
P^{th}	Set of thermal generation types and index, $e \in P^{th}$, $P^{th} \subseteq E$
P^{ht}	Set of heating technologies at the factory layer, $P^{ht} = P^b \cup P^e$
P^b	Set of biomass heating technologies at the factory layer, $P^b \subseteq P^{ht}$
P^e	Set of electricity heating technologies at the factory layer, $P^e \subseteq P^{ht}$
S^{tr}	Set of transformer types and index, $s \in S^{tr}$
V	Set of transmission line types and index, $v \in V$
R^d	Set of distribution line types and index, $r^d \in R^d$
Ω	Set of all scenarios / sampled parameter combinations, $\omega \in \Omega$
Ω^{bw}	Set of valid biomass heating technology-feedstock pairs (b, w)
ε	Sets of electricity generation type-site pairs, (e, j), where $\varepsilon = \varepsilon^{exist} \cup \varepsilon^{planned}$
$SECTOR$	Set of sectors for electricity consumption, $sector \in SECTOR$

Superscript

Symbol	Description
$base$	Relates to base case
bio	Relates to biomass
dem	Relates to demand
$dist$	Relates to distribution lines
e	Relates to electricity
fac	Relates to factory
gen	Relates to generation sites
ht	Relates to heating technology
in	Relates to inlet of a component
$MeOH$	Relates to green methanol
out	Relates to outlet of a component
POC	Relates to POC
st	Relates to storage
$subtrans$	Relates to sub transmission
sup	Relates to supply
$trans$	Relates to transportation
(1)	Relates to first stage optimisation model
(2)	Relates to second stage optimisation model
2037	Relates to the projected year 2037
ω	Relates to Monte Carlo run

Units

Symbol	Description
˚C	Degree Celsius
EJ	Exajoule
EUR	Euro
GJ	Gigajoule
GWh	Gigawatt-hour
kmol	Kilomole
kWh	Kilowatt-hour
MVA	Megavolt-ampere
MW	Megawatt
MWh	Megawatt-hour
NZD	New Zealand dollar
PJ	Petajoule
t	Tonne
TWh	Terawatt-hour
yr	Year

Contributing Publications

The work contributing to this thesis has been published over the past three years in several international conferences and one journal, and both the journal article and the conference proceedings were peer-reviewed.

Journal Article

1. Bertók, B., Chong, D.J.S. & Frits, M. Modelling optimal investment planning for household photovoltaic and battery systems under dynamic electricity market conditions. *Discover Sustainability* **7**, 384 (2026). https://doi.org/10.1007/s43621-026-02683-2.

Conference Proceedings

2. Chong, D. J. S., Walmsley, T., Atkins, M., Bertók, B., Walmsley, M. R. W. 2025. *Aotearoa New Zealand's Energy Future: A Model for Industrial Electrification through Renewable Integration*. Systems Control and Transitions, 4, 1139 to 1144. ESCAPE-35 Conference.
3. Chong, D. J. S., Walmsley, T. G., Atkins, M. J., Bertók, B., Walmsley, M. R. W. 2025. *Low Carbon Industrial Process Heat Scenarios for New Zealand using a P-graph Approach.* PRES 25 Conference.

Table of Contents

Abstract ii
Acknowledgments iv
Nomenclature v
Chapter 1: Introduction 1
1.1 Thesis Aim 3
1.2 Thesis Structure 4
Chapter 2: Literature Review 6
2.1 Introduction 6
2.2 Decarbonising Industrial Process Heat 7
2.3 Green Methanol as a Sensitivity Case 20
2.4 Energy System Modelling Approaches 22
2.5 Existing Energy Systems Modelling 33
2.6 Optimisation and Modelling Frameworks 37
2.7 Conclusions 52
Chapter 3: Methodological Framework 56
3.1 Introduction 56
3.2 Modelling Concept 56
3.3 Section A: Framework Development 60
3.4 Section B: Case Study Workflow and Comparative Analysis of Simulation Modes 62
3.5 Chapter Summary 63
Section A: Energy Modelling Framework Creation 65
Chapter 4: Modular Architecture of Components in the Energy Modelling Framework 66
4.1 Introduction 66
4.2 Methodological Workflow 67
4.3 General Scope Definition 68
4.4 Key Layers and Components 69
4.5 P-graph Modelling Techniques 85
4.6 Modular Components Required for the Energy Model 91
4.7 Chapter Summary 94
Chapter 5: Framework and Mathematical Formulation 95
5.1 Introduction 95
5.2 General Import of Data 96

5.3 Main Model Structure 96
5.4 Coupling P-graph with Python 98
5.5 P-graph-to-Pyomo Port Tool 104
5.6 Pyomo Model Structure 111
5.7 Results Standardisation and Post-Processing 117
5.8 Illustrative Example of the Framework and Modular Operations 119
5.9 Chapter Summary 125
Section B: Case Study Workflow and Comparative Analysis of Simulation Modes 127
Chapter 6: System Definition and Context for New Zealand 128
6.1 Introduction 128
6.2 Scope and Modelling Dimensions 128
6.3 Data Sources and Reconciliation Issues 129
6.4 Process Heat 131
6.5 Electricity Generation 138
6.6 Biomass Resource 143
6.7 Point of Connections (POCs) and Electricity Grid Assets 148
6.8 Green Methanol as a Sensitivity Case 156
6.9 Methods of Geocoding and Distance Calculation 157
6.10 Verification for Model Inputs 162
6.11 Chapter Summary 169
Chapter 7: Multiperiod Optimisation 171
7.1 Introduction 171
7.2 Deterministic Multiperiod Solution Approach 171
7.3 Model Formulation 173
7.4 Constraints 177
7.5 Outputs and Reporting 180
7.6 Sensitivity and Robustness Checks 196
7.7 Limitations and Conclusions 198
Chapter 8: Sequential Optimisation 200
8.1 Introduction 200
8.2 Model Overview 201
8.3 Sequential Optimisation Modelling Framework 206
8.4 Model Formulation 208

8.5 Output and Reporting 211
8.6 Output Structure and Discussion 212
8.7 Limitations and Conclusions 227
Chapter 9: Monte Carlo Uncertainty Analysis 229
9.1 Introduction 229
9.2 Monte Carlo Concept 230
9.3 Temporal Structure and Stochastic Parameters 231
9.4 Model Formulation 236
9.5 Output and Reporting 238
9.6 Chapter Summary 248
Chapter 10: Conclusions and Recommendations for Future Work 249
10.1 Section A: Energy Modelling Framework Creation 250
10.2 Section B: Case Study Workflow and Comparative Analysis of Simulation Modes 250
10.3 Future Work on Energy Modelling Framework 252
10.4 Future Case Study Extensions and Energy Modelling Architecture 254
References 256
Appendix A 279
Appendix B 283
Appendix C 289
Appendix D 295
Appendix E 297
Appendix F 299

List of Figures

Figure 2.1: Hydrogen Ladder 5.0, adapted by the author based on Liebreich (2025). 10
Figure 2.2: Energy End Use Database (EEUD) of New Zealand, compiled by the author based on EECA (2024). 11
Figure 2.3: Temperature ranges of selected processes based on NACE classifications, prepared by the author using data from the European Commission (2008). 12
Figure 2.4: Overview of indirect heating technology selection in the market, adapted from von Thadden del Valle et al. (2024). 13
Figure 2.5: Heat pump cycle, adapted from Koundinya and Seshadri (2022). 14
Figure 2.6: Coefficient of performance (COP) of heat pumps as a function of temperature lift, reproduced with permission from Arpagaus et al. (2018). 16
Figure 2.7: Schematic diagram of a coal-fired or biomass-fired grate boiler design, adapted from Huang et al. (2023). 17
Figure 2.8: Block diagram of a biomass-to-methanol plant, adapted by the author from Poluzzi, Guandalini and Romano (2022). 21
Figure 2.9: A depiction of an energy system model in a directed graph approach and a typical backend architecture on components representing specific respective energy technologies adapted from Hoffmann et al. (2024). 26
Figure 2.10: Overview of methods for reducing temporal complexity in discretised time space, adapted from Kotzur et al. (2021). 30
Figure 2.11: Mathematical program variations based on investment cost formulations, adapted from Kotzur et al. (2021). 33
Figure 2.12: P-graph representation of a process network problem. 39
Figure 2.13: Illustration of a simplex method application by solving a linear programming problem with x and y variables. 42
Figure 2.14: The convex and non-convex problems when dealing with a non-linearity problem adapted from Zhang and Xie (2023). 44
Figure 2.15: Infeasible integer values from relaxed MILP adapted from Grossman (2021). 46
Figure 2.16: Superstructure model with disjunctive operation, adapted from Grossmann (2021). 47
Figure 2.17: Comparison of search space between convex hull and Big-M formulation. 49
Figure 2.18: Schematic of the Benders decomposition method, adapted from Rahmaniani et al. (2017). 50
Figure 3.1: Venn diagram illustrating the relationship between the three elements. 57
Figure 3.2: Overall methodological framework for the thesis. 59
Figure 4.1: Overview of the methodological workflow for Section A. 67
Figure 4.2: Superstructure of the proposed energy modelling framework. 70
Figure 4.3: Monthly operational hours across different industrial sectors in New Zealand. 74
Figure 4.4: Normalised generalised operating efficiency ratio, prepared by the author using data from Forschungsgesellschaft für Energiewirtschaft (FfE) (2021). 74
Figure 4.5: Average per capita electricity consumption by region, prepared by the author using data from (Electricity Authority, 2025). 79

Figure 4.6: Average number of people per ICP, prepared by the author using data from (Electricity Authority, 2025). 80
Figure 4.7: Projected EV electricity demand (TWh yr^{-1}), based on data from MBIE (2024b). 81
Figure 4.8: Conceptual representation of determining non-industrial process heat demand at each POC node. 82
Figure 4.9: Net production costs of green methanol at different plant scales based on electrolyser-rated power, based on data from Pratschner et al. (2023), assuming 1 EUR equals 2 NZD. 84
Figure 4.10: Illustration of time-step modelling applied in P-graph. 86
Figure 4.11: P-graph representation of a proportional capacity expansion problem. 87
Figure 4.12: P-graph solution illustrating proportional capacity sizing. 88
Figure 4.13: P-graph representation of an integer-based capacity expansion problem. 89
Figure 4.14: P-graph solution illustrating integer-based capacity sizing. 90
Figure 4.15: P-graph solution showing integer capacity expansion at time step $t = 2$. 90
Figure 4.16: Modular components implemented in the energy modelling framework. 91
Figure 4.17: Example of component layer data structure, showing descriptive, spatial, and functional attributes, along with metadata and annotations. 92
Figure 4.18: Modular component for the factory layer with |T|=1. 93
Figure 5.1: Model overview including required prompts, configurations and exemplary outputs. 96
Figure 5.2: Adapted P-graph and Python workflow based on Teng et al. (2022). 98
Figure 5.3: Standard P-graph workflow for process network synthesis. 98
Figure 5.4: P-graph and Python workflow extended from the conventional version. 99
Figure 5.5: Problem formulation in Pyomo and the transformation process to a solvable MILP. 100
Figure 5.6: Comparison of output files generated by P-graph Python (left) and P-graph Studio (right). 102
Figure 5.7: Implemented output sheets and their attributes in the Excel spreadsheet. 103
Figure 5.8: Proposed integrated P-graph and Pyomo workflow. 104
Figure 5.9: Options for modelling pipelines within the solving framework. 105
Figure 5.10: Modelling pipeline for the Pyomo-to-P-graph framework. 105
Figure 5.11: Comparison of input file formatting for P-graph Python (left) and P-graph Studio (right). 106
Figure 5.12: Regular expression rule enforced in the Flow Section, where (a) shows the general regular expression rule and (b) shows the example and grouping names. 110
Figure 5.13: Regular expression rule enforced in the Mutual Exclusion Section, where (a) shows the general regular expression rule and (b) shows the example and grouping names. 111
Figure 5.14: P-graph structure represented using an equation-based approach. 113
Figure 5.15: Conceptual design and rationale for presenting post-processed outputs with descriptive attributes. 117
Figure 5.16: Identifier for the energy input to the respective heating technologies. 119
Figure 5.17: Identifier for the capacity sizing for the respective heating technologies. 119
Figure 5.18: Ranking of all feasible solutions by total cost. 121

Figure 5.19: Single time-step system with one generation, factory, biomass, and methanol unit. 122
Figure 5.20: Modelling summary showing the relationship between the number of constraints and the number of time steps. 125
Figure 6.1: Industrial process heat demand in New Zealand, prepared by the author using data from EECA (2024). 133
Figure 6.2: (a) Decarbonisable industrial process heat in New Zealand, compiled from RETA and the Heat Plant Database; (b) Spatial distribution of decarbonisable process heat temperature levels, based on Table 4.2 and European Commission (2008), supplemented by referenced and partly confidential industrial data 134
Figure 6.3: Sectoral distribution of decarbonisable process heat temperature levels in New Zealand, based on Table 4.2 and European Commission (2008), with additional referenced and partly confidential industrial data 134
Figure 6.4: Decarbonisable process heat demand ranked in descending order with cumulative distribution 135
Figure 6.5: Compiled process heat dataset represented in a Venn diagram. 136
Figure 6.6: Number of coal boilers with retrofitting potential by industrial sector in the model. 137
Figure 6.7: Spatial distribution of coal boilers with retrofitting potential by industrial sector in the model 137
Figure 6.8: Installed capacities by plant type (excluding hydroelectricity) (MBIE, 2024b). 139
Figure 6.9: Comparison of peak demand ratio and load ratio from Jan 2010 to Sep 2022 (Transpower New Zealand, 2023b) 140
Figure 6.10: Top 20 daily peak electricity demand values from January 2010 to September 2022 (Transpower New Zealand, 2023b) 140
Figure 6.11: Spatial distribution of intermittent generations in New Zealand. 142
Figure 6.12: Agricultural and planted forest cover data from the New Zealand Land Use Map (Ministry for the Environment NZ, 2025) and locations of wood pellet suppliers. 144
Figure 6.13: Spatial distribution of biomass resource types by gross volume (Hall, 2022) 145
Figure 6.14: Bar chart on biomass resource type by gross volume by region (Hall, 2022). 145
Figure 6.15: Biomass location nodes in New Zealand, identified using geospatial analysis. 147
Figure 6.16: The concept of spot electricity market pricing in New Zealand. 148
Figure 6.17: Wholesale electricity prices (logarithmic-scaled) from 2021 to 2025 at hourly resolution, based on data from the Electricity Authority (2025). 149
Figure 6.18: Electricity consumption by sector based on data from MBIE (2024b). 150
Figure 6.19: Adjusted price ratio across GXPs. 152
Figure 6.20: Electricity consumption in 2022 by sector (EV, industrial, residential, and commercial), based on model assumptions. 154
Figure 6.21: Spatial representation of POC and transmission layers, including HVDC infrastructure, across New Zealand with data from Transpower New Zealand (2025b). 155
Figure 6.22: Candidate locations for green methanol production, classified by proposed generation type 157
Figure 6.23: Raster output from the Distance Accumulation tool in ArcPy. 159

Figure 6.24: Counts of biomass-to-generation connections available as model inputs generated from the Distance Accumulation tool. 160
Figure 6.25: Counts of biomass-to-factory connections available as model inputs generated from the Distance Accumulation tool. 160
Figure 6.26: Comparison of actual daily wind data from MetService (2025) and simulated wind data from Renewables.ninja from 2009 to 2016 overlaid in the same graphical plot. 163
Figure 6.27: Actual aggregated monthly wind data from MetService (2025) and simulated wind data from Renewables.ninja from 2009 to 2016 overlaid in the same graphical plot. 163
Figure 6.28: Actual average solar irradiance data from Solarview (NIWA, 2025) and simulated data from Renewables.ninja in weeks overlaying in the same graphical plot 164
Figure 6.29: Actual average solar irradiance data: (a) seasonal variation shown in a line plot, and (b) hourly distribution by month presented as a heatmap. 164
Figure 6.30: Monthly electricity demand profiles in GWh for (a) Penrose, (b) Edendale, (c) Glenbrook, and (d) Paraparaumu, showing total actual electricity demand (black solid), estimated total electricity demand (black dotted), estimated industrial demand (orange dotted), and residential and commercial demand (blue dotted). 166
Figure 6.31: Comparison of reference year vs projected year for process heat optimisation. 167
Figure 6.32: Projected sectoral electricity demand by region in 2037, excluding electrification of decarbonisable industrial process heat. 168
Figure 7.1: Superstructure model of sets and indices in the mathematical formulation. 174
Figure 7.2: Regional distribution of decarbonisable process heat by input fuel to heating technologies. 181
Figure 7.3: Fuel-switching technologies by input fuel in the Auckland region. 182
Figure 7.4: Decarbonisable process heat by sector and input fuel for heating technologies. 183
Figure 7.5: Decarbonisable process heat by sector with end-use demand for heating technologies. 183
Figure 7.6: The spatial distribution of retrofitted coal boilers. 184
Figure 7.7: The spatial distribution of the biomass connections to factories. 185
Figure 7.8: Number of biomass to factory connections established in the output results. 185
Figure 7.9: Biomass resources utilised and upper limit by region. 186
Figure 7.10: Figure of biomass type consumed by region for process heat industry. 186
Figure 7.11: Example biomass source (ID 81) and its connections. 187
Figure 7.12: Total biomass flow from ID 81 to each destination ID aggregated over the full year as the sum of monthly flows. 187
Figure 7.13: Number of green methanol feasibility sites by varying selling price. 188
Figure 7.14: Possible green methanol sites when selling price is set as NZD 950 NZD t^{-1} 189
Figure 7.15: Model results in GWh yr^{-1} for the Fonterra Stirling site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%. 191
Figure 7.16: Model results in GWh yr^{-1} for the Fonterra Whareroa site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%. 191
Figure 7.17: Model results in GWh yr^{-1} for the Fonterra Clandeboye site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%. 192

Figure 7.18: Model results in GWh yr^{-1} for the Fonterra Hautapu site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 400%. 192
Figure 7.19: Model results in GWh yr^{-1} for the Fonterra Waitoa site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%. 193
Figure 7.20: Model results in GWh yr^{-1} for the Fonterra Edendale site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 100% and electricity price for the Edendale site is halved. 194
Figure 7.21: (a) Optimisation model results in GWh yr^{-1} for the Fonterra Edendale site under the base scenario; (b) temperature-based demand requirements of the main industrial processes at the site. 194
Figure 7.22: (a) Optimisation model results in GWh yr^{-1} for the NZ Sugar site under the base scenario; (b) temperature-based demand requirements of the main industrial processes at the site. 195
Figure 7.23: Model results in GWh yr^{-1} for the Open Country Waharoa site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%. .. 195
Figure 7.24: Sensitivity analysis of electrification uptake in industrial process heating (excluding wood processing and pulp and paper sectors). 197
Figure 8.1: Illustrative example showing how the factory layer is simplified between stages in the multistage optimisation approach. 204
Figure 8.2: Illustration of a single time-step energy system with grid reinforcement, consisting of one generation unit, two POCs, one methanol plant, and one industrial factory. 205
Figure 8.3: Model architecture of the two-stage optimisation model. 207
Figure 8.4: Installed new generation capacity by technology and region. 215
Figure 8.5: The generation portfolio including existing and proposed assets implemented in the model solution. 215
Figure 8.6: Sub transmission network expansion from newly installed generation to POC nodes. 216
Figure 8.7: Newly reinforced and existing transmission network assets. 218
Figure 8.8: Distribution lines generated from the electrified process heat in industries to the POC nodes. 219
Figure 8.9: Capital cost breakdown of infrastructure upgrades. 220
Figure 8.10: Fuel-switching outcomes identified in the optimised solution by region. 223
Figure 8.11: Fuel-switching outcomes identified in the optimised solution by sector. 223
Figure 8.12: Regional spatial distribution of biomass and factory connections established in the optimised solution. 224
Figure 8.13: Biomass resources used for process heat in industries by region. 224
Figure 8.14: (a) Fully electrified scenario-based on industrial end-use demand in TWh yr^{-1}. (b) Base case model in TWh yr^{-1}. (c) Distribution of heating demand across temperature bands TWh yr^{-1}. 226
Figure 8.15: Spatial distribution of new proposed generation in MW, including retrofitting thermal and cogeneration generation and commissioning bioenergy capacity for the full electrification scenario. 226
Figure 9.1: Illustrative example of convergence band for each Monte Carlo run. 231

Figure 9.2: Hot water storage design within the P-graph framework, $T = 1$. 237
Figure 9.3: Monte Carlo probability density function plot for total system cost. 238
Figure 9.4: Cumulative running mean in a convergence plot. 239
Figure 9.5: Plot of ASCBR consecutive count of random variable zj. 239
Figure 9.6: Plot of running standard deviation across iteration runs. 240
Figure 9.7: System-level results on heating technologies and thermal storage options. 242
Figure 9.8: System-level results on biomass resource utilisation and availability. 243
Figure 9.9: Individual biomass nodal ID=2 level results on biomass resource utilisation and availability. 245
Figure 9.10: Individual industrial site ID=8 level perspective on heating technologies and thermal storage potential. 246
Figure 9.11: Individual biomass nodal ID=2 level results connection to 12 factory sites. 247

List of Tables

Table 2.1: Overview of carbon abatement for industrial processes (Olsson and Schipfer, 2021). . 7
Table 2.2: High-temperature industrial processes. 9
Table 2.3: List of new heat pump technologies and their specifications (Zühlsdorf *et al.*, 2023). 15
Table 2.4: Temperature levels of industrial heat across literature sources. 19
Table 2.5: Summary of technological aggregation from literature sources. 31
Table 2.6: Specifications on constructing an energy modelling framework (Lahijani, Protheroe and Gschwendtner, 2025). 35
Table 2.7: Strategies of linearisation methods for non-linear and non-continuous constraints which exist in energy system models. 43
Table 4.1: Heating technologies introduced within the framework for indirect heating. 71
Table 4.2: Default process heat temperature shares by process category, based on the European Commission (2008) and additional referenced and partly confidential industrial data. 71
Table 4.3: Levelised cost for new generation. 75
Table 4.4: Classification of boiler types by biomass feedstock. 76
Table 4.5: Conversion and energy basis assumptions for green methanol representation. 84
Table 5.1: Regular expression applied to parse default parameter assignments, including examples of captured groups and parsed results. 107
Table 5.2: General regular expression rules applied to input and output files for data parameterisation. 107
Table 5.3: General regular expression rule in the Material Section. 108
Table 5.4: General regular expression rule in the Operating Units Section. 109
Table 5.5: List of parameters and bound constraints for each component. 112
Table 5.6: Model components, specifications, and associated variable types. 112
Table 5.7: Optimisation solvers tested by author in the Pyomo environment, including supported problem classes, licensing, and implementation details. 115
Table 5.8: Naming convention used to extract attributes from reference layer identifiers from the dataset. 118
Table 5.9: Installed heating technology capacities for the best optimal solution structure. 123
Table 5.10: Summary of selected feasible heating technology configurations and associated system costs. 124
Table 5.11: Model size and computational characteristics for different numbers of time steps. 125
Table 6.1: Comparison of sectoral energy fuel consumption (TWh) between compiled results and national datasets from MBIE (2024b) and EECA (2024) for the year 2022. 132
Table 6.2: Summary of industrial sites not included in the model and their associated process heat demand for decarbonisation (GWh yr^{-1}). 136
Table 6.3: Example industrial facilities used as model inputs for the nationwide process heat decarbonisation analysis. 138
Table 6.4: Example input parameters for electricity generation sites in the model. 142
Table 6.5: Cost specifications by regions and net calorific value. 143
Table 6.6: Recoverability factors based on different estimations by Hall (2022). 145

Table 6.7: Example biomass center inputs. 147
Table 6.8: Estimated cost components fractions for large energy users (Electricity Offers NZ, 2025). 151
Table 6.9: Electricity price ranges for key GXPs implemented in the model. 152
Table 6.10: Example POC input specifications, including technical and demand-related parameters for selected nodes 155
Table 6.11: Biomass-generation sites from the iterative Data Accumulation function. 159
Table 6.12: Biomass-factory sites from iterative Data Accumulation function. 161
Table 6.13: Distance estimation within POCs and the Distribution Level. 162
Table 6.14: MBIE Energy balance table based on ratio (MBIE, 2024b). 165
Table 7.1: Summary of optimisation model size and computational performance. 180
Table 7.2: Summary of company sites identified for comparison. 190
Table 7.3: Recoverability factors based on different estimations by Scion for sensitivity analysis (Hall, 2022). 196
Table 8.1: Summary of the sectoral electricity demand and average sales electricity price for base value (MBIE, 2024a). 202
Table 8.2: Summary of decision scope and assumptions between the Stage 1 and Stage 2 optimisation framework. 203
Table 8.3: Summary of optimisation model size and computational performance for the grid reinforcement model. 211
Table 8.4: Uncertainty analysis based on grid infrastructure cost. 212
Table 8.5: Anticipated transmission network upgrades relevant to modelling results. 217
Table 8.6: Comparison of fuel-switching results by end-use demand between Chapter 7 and the sequential grid model. 222
Table 9.1: Stochastic parameters implemented in the Monte Carlo uncertainty analysis. 235
Table 9.2: Factory specifications connected by biomass node ID=2. 244

Chapter 1: Introduction

The industrial sector contributed more than 38% of global CO_2 emissions in 2022 (IEA, 2024). Projections by the International Energy Agency (IEA) indicate that, in a 2050 Net Zero scenario, around 90% of industrial emissions must be reduced to limit global warming to 1.5°C above preindustrial levels, as outlined in the Paris Agreement. A large share of industrial emissions arises from the use of fossil fuels for process heating. These fuels have historically been preferred because of their affordability and abundance during the twentieth century. However, growing concerns regarding climate change and resource constraints have intensified the need to decarbonise industrial heat generation. As a result, the transition of industrial process heat systems away from fossil fuels has become an essential component of global decarbonisation strategies.

Despite growing electrification in transport and some manufacturing applications, fossil fuels remain the dominant energy source for industrial process heating. Globally, the share of renewable fuels in industrial heat is expected to increase only slowly, from 12% in 2022 to 15% in 2028 (IEA, 2024). At the same time, industrial heat demand is projected to increase by roughly 16% over the same period. Electrification has emerged as a promising pathway for decarbonisation, with its share expected to rise from about 4% to 11% within the industrial heat sector (IEA, 2024). Electrified heating technologies such as heat pumps and electric boilers offer several advantages, such as operational simplicity, and reduce on-site emissions when supported by renewable electricity generation.

Energy systems that increasingly rely on renewable sources for electricity generation make these technologies particularly appealing. Under such conditions, electrified heating technologies can significantly reduce the carbon intensity of industrial heat production. However, electrification alone cannot universally resolve the challenge of industrial process heat decarbonisation. Large-scale electrification has to be supported by sufficient generation capacity to maintain security of supply. Without adequate expansion of electricity generation infrastructure, increased demand may create shortages during peak periods.

In some cases, a centralised electricity supply may not be an economic option for industrial process heat users located in rural regions. Bioenergy is another potential pathway to phase out fossil fuels. Global bioenergy use is projected to increase to 4.2 EJ in the IEA 2023 to 2028 Net Zero scenario, indicating that biomass is expected to play a prominent, if not dominant, role in decarbonising industrial process heat (IEA Bioenergy, 2024). A key advantage of biomass is the potential to retrofit existing boiler systems that currently use solid fossil fuels, rather than replacing them entirely, which can reduce lifecycle costs. However, the proximity between industrial sites and biomass supply centres is important, as longer transport distances increase logistics costs and therefore fuel costs.

To optimise an overall energy system at this level of detail, accounting for major industrial sites, electricity generation plants, biomass locations, and grid networks at the site level, a modelling tool must represent all these features within a single framework.

Conventional macroscale energy system tools tend to be highly specialised and operate at a relatively generic, high-level abstraction. Many focus on a single primary function such as transport planning, district heating, electricity market pricing, or emissions analysis, which do not capture any cross-sectoral complexity. Tools that span multiple technologies and system layers often rely on aggregated representations of demand in order to simplify the system into a single component block. Conversely, macroscale energy system models with high spatial and temporal resolution frequently lack detailed process-level representation within individual components.

A key limitation of existing modelling tools is the inability to represent process heat decarbonisation at a spatial level. Unlike hydrogen, electricity, and other energy commodities, process heat has both quantitative and qualitative characteristics. Process heat demand can be differentiated by temperature levels, and heat can be cascaded from higher to lower temperature uses, depending on site requirements. As a result, aggregation of heat demand or temperature requirement is challenging, and fuel-switching options are highly site-specific. Many existing tools cannot easily support customised modular operating units. Consequently, modellers often simplify the problem, shifting the research objective away from identifying cost-optimal structural system designs and towards higher-level projections of renewable energy mixes.

Generic optimisation environments are often used to address specific energy modelling problems because algebraic modelling tools such as Python-based frameworks, GAMS, and Julia-based tools can be powerful. Solver implementations in these environments are mature and capable of handling large-scale models. However, repeatability and reproducibility are often limited because models are implemented as code and are frequently used only within the developing institution. Such implementations often involve an extensive set of equations, which can be difficult for readers or other modellers to interpret and replicate.

An alternative generic approach is Process Network Synthesis (PNS). PNS represents process flows using bipartite graphs and applies a combinatorial approach to identify feasible pathways by linking inputs to products. Optimisation has gained attention for this method due to its applicability to diverse problems, including process engineering, evacuation routing, transport planning, and energy allocation. A key advantage of Process Network Synthesis (PNS) is its transparency: connections between components are explicit and easy to interpret. PNS is the overarching concept and approach, which is implemented using P-graph, a graph-theoretic framework for modelling and optimising process systems. In addition, PNS is supported by an open-source interface, P-graph Studio, which lowers the barrier for engineers and energy modellers to develop models without requiring advanced coding skills or deep expertise in mathematical programming.

However, P-graph faces challenges for large-scale problems, as visual representations become impractical due to memory and display constraints. Although a Python-based library has been developed to support automation and iterative model construction, the P-graph's built-in Accelerated Branch and Bound solver, which applies a branch and bound concept, is more suited to small-to-medium sized problems because the number of subproblems in its branch and bound

search grows exponentially (Eichfelder *et al.*, 2021). Currently, the P-graph Python environment also offers limited solver options.

These limitations, which include restricted solver availability in P-graph Python, the difficulty of developing and interpreting large-scale mathematical programming models, and the need for intensive manual work when assigning and connecting nodes in the interface, represent key bottlenecks in developing energy system models that can simultaneously address grid expansion, process heat constraints, and logistics.

Developing an energy modelling tool with such a wide scope therefore requires a strong foundation in process systems engineering, optimisation, energy markets, mathematical programming, energy transformation, supply chain logistics, geospatial modelling, and software development. At the same time, usability must also be prioritised. This includes enabling users to implement generalised disjunctive programming (GDP) without manually writing formulations, optional preprocessing tools helpful for exogenous input data generation, supporting flexible modular unit definitions for iterative modelling, selecting the number of time periods, offering optimisation using both deterministic parameters and randomly sampled Monte Carlo approaches, and allowing export of the model structure into an interface for inspection and further development.

1.1 Thesis Aim

The aim of the research reported in this thesis is to evaluate nationwide fuel-switching pathways for medium-temperature industrial process heat at the site level. Process heat temperature profiles are limited to up to 500°C, reflecting the range in which economically viable technologies such as heat pumps, electric boilers, and biomass boilers can be applied. A focus is placed on grid expansion and electricity generation, as these strongly influence overall fuel costs.

The thesis develops a spatial energy modelling framework that combines P-graph superstructure generation with automated translation to a Pyomo-based GDP optimisation model, enabling large-scale structural optimisation and flexible simulation modes. The spatially explicit energy modelling framework is developed to determine the least-cost national system design for supplying renewable energy to major industrial sites.

The framework is designed to be generalisable, meaning it can be applied in any national or regional context provided that the input data are specified consistently. It also incorporates multiperiod modelling and several solution approaches, including single deterministic optimisation, sequential deterministic optimisation, and Monte Carlo-based simulation. Scenario-based analysis is used to assess how four-hourly time steps and seasonal variation can influence system configurations.

The framework integrates key components required for a comprehensive assessment. This includes electricity generation and the transmission grid, so as to evaluate grid and substation upgrades driven by process heat electrification. The feasibility of retrofitting coal boilers to

biomass boilers is also included, as it captures the techno-economic implications of boiler upgrades compared with full equipment replacement.

As a proxy for the broader transition required to meet decarbonisation targets, the framework includes sensitivity analyses for emerging green industries that may compete with industrial process heat loads for limited renewable energy resources. The main output of this research is a generalised modelling framework composed of modular and modifiable components. With minimal reliance on explicit equation-based formulation in the user workflow, it is intended to enable energy modellers to develop large-scale, high spatial resolution energy models more efficiently.

The framework is implemented as a transparent, replicable, and adaptable Python-based energy modelling tool, designed for application in other countries or regions and for supporting comparative analysis of decarbonisation pathways. Model outputs are generated in a structured spreadsheet format and then post-processed for visualisation and analysis to facilitate efficient interpretation and integration.

In summary, to achieve the aim, the thesis addresses the following research questions:

I. What range of renewable energy types and heating technologies are required to phase out fossil-based fuels in industrial process heat while remaining within resource constraints?
II. What upgrades to the electricity grid are necessary to accommodate electricity use in the projected year and to support increased fuel-switching for industrial process heat and fluctuations of heating demand within factories?
III. To what extent does accounting for seasonal and monthly temporal variation, relative to a single annual snapshot, influence the least-cost national system design for supplying a renewable energy solution to all major industrial sites?

The primary contribution of this thesis is a generalisable methodology for coupling Process Network Synthesis, via P-graph with algebraic disjunctive programming via Pyomo.GDP. This addresses a recognised gap in process systems engineering in linking P-graph's structural representation to scalable mathematical programming solvers. The secondary methodological contribution is in modularising spatially explicit energy system representation and multi-stage optimisation, alongside empirical and policy findings from a New Zealand case study. These contributions are detailed in Section 2.7.1

1.2 Thesis Structure

This thesis is organised into several chapters that progressively develop the modelling framework and demonstrate its application to industrial process heat decarbonisation with a New Zealand case study.

Chapter 2 reviews the relevant literature on industrial process heat decarbonisation, energy system modelling approaches, and optimisation methods used in large-scale energy system studies. The chapter highlights the limitations of existing modelling tools in representing industrial

process heat at a spatially detailed level. In addition, mathematical programming methods are discussed to emphasise their importance in the backend formulation of the optimisation problem.

Chapter 3 presents the methodological framework that defines the core contribution of this thesis. It introduces the modelling approach that integrates PNS with algebraic modelling environments to support the development of large-scale energy system models. The second half of the chapter outlines the case study workflow and presents a comparative analysis of simulation modes, which establishes the foundation for the subsequent chapters.

Chapter 4 presents the structural architecture of the energy modelling framework. The chapter explains how modular units and system connections are defined within the framework to allow flexible construction of energy system models.

Chapter 5 presents the computational and mathematical implementation of the framework. It explains the methodology on how P-graph superstructure is translated into a Pyomo-based optimisation model. The chapter also includes data import, translation logic, solver integration, and standardised outputs.

Chapter 6 builds upon the energy modelling framework by employing a case study of process heat decarbonisation strategies in New Zealand industries. Optional preprocessing tools for generating exogenous model inputs are also introduced. Scenario-based analysis is used to evaluate process heat technology configurations across the modelled system.

Chapter 7 presents the results of the deterministic multiperiod optimisation model. The chapter evaluates the least-cost configuration of technologies for industrial process heat decarbonisation under temporal variation in demand and resource availability.

Chapter 8 extends the analysis through a sequential deterministic optimisation approach that considers staged system development and grid reinforcement requirements associated with large-scale electrification.

Chapter 9 applies Monte Carlo-based optimisation to the multiperiod model to assess how uncertainty affects system configuration. The analysis focuses on uncertainty in fuel-related commodities and parameters, which include electricity price, biomass availability, and biomass cost.

Finally, Chapter 10 summarises the major findings, draws overall conclusions, and provides recommendations for future research directions.

Chapter 2: Literature Review

2.1 Introduction

Industrial decarbonisation is pivotal to achieving net-zero emission targets. As a result, modern industries are increasingly subject to government policies and carbon emission pricings that require them to pursue alternative pathways for carbon abatement, efficiency improvement, and fuel switching. However, both policymakers and private entities face significant inertia in adopting new technologies, as technology selection can have system-wide implications and must be compatible with existing energy infrastructure such as electricity grids and fuel pipelines. These pressures are not confined to the plant boundary. Decisions about industrial decarbonisation increasingly interact with wider electricity systems, biomass supply chains, network constraints, and spatially uneven resource availability. As a result, industrial process heat decarbonisation is not only a technology choice problem but also a systems integration problem.

Energy models play a central role in addressing this challenge by providing forward-looking scenario analysis and identifying minimum cost pathways that support informed decision-making. Despite the wide availability of energy modelling tools, many existing models operate at low spatial resolution. Even when spatial detail is included, models are often transmission-centric, focusing on networks rather than on factories and energy demand.

Energy loads at the distribution level, such as industrial sites, can be scaled, while on-site subsystems can be simulated. Identifying energy pathways is important from the demand response perspective, where spatial proximity and existing infrastructure capacity remain important considerations for the energy transition. However, this class of energy models remains underdeveloped due to the complexity of the backend, mathematical formulation, and reusability for a different case study. The limitations above create a gap between plant-level realism and system-level optimisation. This gap motivates the present study. It focuses on industrial process heat as a distinct modelling problem that requires explicit treatment of heat quality, technology compatibility, spatial context and infrastructural constraints. The review narrows the broad fields of energy system modelling, process optimisation and industrial decarbonisation into the specific body of literature most relevant to the aim. In this chapter, literature is reviewed across five interconnected themes.

First, the industrial process heat is reviewed with particular attention to the characteristics of heating demand and decarbonisation pathways. Operating temperature is a key determinant of technology selection. Therefore, process heat demand is analysed based on temperature classification. Second, the evolution of energy system modelling is examined to situate the thesis within the modelling context. This section reviews the fundamentals of energy modelling, including model dimensions and complexity, to clarify how scope, time, and mathematical formulation affect tractability and usefulness. Third, existing modelling tools and generic optimisation environments are critically reviewed to justify the relevance of developing a customised energy modelling framework. Fourth, mathematical programming methods and

solution algorithms are studied before outlining the key research gaps which exist in the literature which will be answered in the thesis.

2.2 Decarbonising Industrial Process Heat

Industrial process heat is highly reliant on fossil fuels. Within a plant-scale boundary, apart from heat insulation, systematic process integration is the most energy-efficient method to reduce energy demand. A key contribution of process integration is the concept of pinch analysis. Heat recovery pinch, or pinch analysis, was famously developed and made widespread by Linnhoff and Flower (1978). Heat recovery pinch has been applied to the heat integration between streams within the process plant and focused on sizing heat exchanger networks to minimise total system costs (Knorr *et al.*, 2025). The concept soon evolved into wider applications, including an extension of heat recovery to include total site integration (Liew *et al.*, 2017). In addition to these energy efficiency measures; sector coupling is a viable pathway to substitute existing fossil fuels with renewable energy sources (Ramsebner *et al.*, 2021). Electrification of heat and replacing fossil fuel with biomass plays a significant role in the process heat industry.

A broad body of literature has emerged on technological developments for decarbonising process heat (Mugarza *et al.*, 2024). To assess large-scale energy models, plant-scale process improvements such as material substitution, process pathway changes, or heat recovery are often overlooked due to the wider system scope. Although process- and product-related emissions contribute to carbon abatement, they are typically treated explicitly within their own sub-boundaries (Malehmirchegini and Chapman, 2025). In contrast, large-scale energy models generally focus on fuel-switching pathways based on spatial proximity and on the quantity and quality of heat required to be delivered. As a result, this modelling approach does not account for process or product changes within the energy system optimisation framework. Table 2.1 outlines the overview of emission abatement options from three different approaches.

Heating technologies can be divided into direct heating and indirect heating. Direct heating applies heat to the process directly, whereas indirect heating refers to the method of transferring the heat source to an intermediate medium rather than to the target object directly. In direct heat applications, a solid material is often heated to a higher temperature using an oven, kiln, burner, or reformer (Reid *et al.*, 2025). The industrial plants using such heating technologies are typically concentrated in a few specific sectors, which are primarily involved in mineral and metal processing (Reid *et al.*, 2025). Fuel-switching is possible, but it is highly case-specific, as it depends on the processing pathway and, in some cases, the fuel is also used as a feedstock. Table 2.2 shows the example of industries where temperature is high for certain processes.

Table 2.1: Overview of carbon abatement for industrial processes (Olsson and Schipfer, 2021).

Fuel Use	**Process Emissions**	**Product Emissions**
Fuel-switching to Biomass Heat (Rissman *et al.*, 2020; Beccarello and Di Foggia, 2023)	CCS (Rissman *et al.*, 2020; Beccarello and Di Foggia, 2023)	Recycling (Rissman *et al.*, 2020)

Table 2.1 (continued): Overview of carbon abatement for industrial processes (Olsson and Schipfer, 2021).

Fuel efficiency increase (Rissman *et al.*, 2020)	Industrial synergy (Wadström, Johansson and Wallén, 2021)	Material Substitution (Tollefson, 2017)
Waste Heat (Rissman *et al.*, 2020)	Process Efficiency (Rubinsin *et al.*, 2024)	Improved product longevity (Cao *et al.*, 2017)
Electrification (Beccarello and Di Foggia, 2023)	New Pathway (Shahabuddin, Brooks and Rhamdhani, 2023)	Intensive product use (Fremstad, Underwood and Zahran, 2018)
Green fuels (Beccarello and Di Foggia, 2023)	Pinch Technology (Walden, Wellig and Stathopoulos, 2023)	Material efficiency (Rissman *et al.*, 2020)
Solar thermal heat (Prestigiacomo *et al.*, 2024)		
Geothermal/ Nuclear heat (Rissman *et al.*, 2020)		

As shown in the table, the heating processes are often heavily reliant on the fuel type. The heating fuel is not only used to achieve the desired operating temperature but is also used as the feedstock for the associated process. For example, in the conventional ammonia and methanol production sector, syngas, which is the source of hydrogen in their processes, is mainly sourced from the Steam Methane Reforming (SMR) reaction, which has an operating range of 800˚C (Kibria, McManus and Bhattacharya, 2023). On the other hand, Haber Bosch process and methanol synthesis process have operating temperatures of 450˚C and 300˚C respectively. Therefore, studies on material substitution using the Power-to-Fuel concept have been conducted to produce electrolysed hydrogen, which is the key ingredient for ammonia and methanol production.

Alternate processing pathways have to be implemented for contemporary steelmaking processes (Reid *et al.*, 2025). Recent modern blast furnaces inject natural gas and oil to replace a fraction of metallurgical coke which is used to turn iron ore into molten iron. The coking process removes impurities and oxygen from iron oxide, which is important in the steelmaking process. Research into the use of hydrogen gas as a reducing agent has been conducted (Montanaro *et al.*, 2023). While the technology has been demonstrated at pilot scale and is technically viable, it remains too nascent for large-scale deployment. This is primarily due to the significant technological modifications required to existing steelmaking operations to implement direct hydrogen reduction of iron. In addition, associated fuel-switching for process heating would also be necessary to achieve the high temperatures required for melting iron.

Table 2.2: High-temperature industrial processes.

Industry	Feedstock	Heating Technology	Temp	Description
Ammonia (Kibria, McManus and Bhattacharya, 2023)	72% natural gas, 26% oil	Furnace, Reformer	800˚C	SMR and coal gasification processes have to be replaced with a new electrolysis pathway for hydrogen production. Studies of biomass to ammonia have been conducted too (Rossetti *et al.*, 2023).
Cement (Clark *et al.*, 2024)	57% coal, 24% natural gas	Kilns	1400˚C	The calcinated lime enters the kiln to be transformed into clinker. High-carbon fuel is currently replaced with natural gas. Studies have assessed the potential of hydrogen fuel (Rissman *et al.*, 2020).
Steel (Manning and Fruehan, 2001)	74% coal, 10% natural gas (IEA, 2020)	Furnaces	2000˚C	The blast furnace involving the process of reducing iron ore to molten iron is an energy-intensive process. The coking process from metallurgical coal has an adverse effect on the environment due to hydrocarbon emissions.
Methanol (Dalena *et al.*, 2018)	90% Natural gas	Furnaces	800˚C	Despite the methanol synthesis reaction to be around 200-300˚C, paired with pressures of 50-100 bar, the syngas production from SMR exceeds 800˚C.

Within the non-metallic minerals, cement produces a large carbon footprint from processing the raw materials which are made up of limestone, clay, shale and other additives. At the initial stages, the mixture undergoes physical processes and forms a raw meal. Subsequently, the raw meal has to be heated to 1400˚C which involves the chemical decomposition of calcium carbonate into calcium oxide, which emits carbon dioxide (Ishak and Hashim, 2015; Barbhuiya, Bhusan Das and Adak, 2024). Special binders or kilns have to be designed to replace conventional coal and natural gas-powered ones which not only replace the fossil fuels but also exhibit carbon capture capabilities for eliminating the carbon dioxide by-product from the decomposition process (Oberschelp, Pfister and Hellweg, 2023).

Hydrogen is primarily considered for high-temperature industrial process heat applications where direct electrification is not technically feasible. It can be used either as a combustion fuel or as a feedstock for synthetic fuels, enabling the substitution of fossil fuels in hard-to-abate sectors such as cement, steel, and chemical production. However, hydrogen-based pathways are associated with significant efficiency losses due to energy conversion processes (Begemann *et al.*, 2021), as well as high production costs, infrastructure requirements (Badakhsh and Mothilal Bhagavathy,

2024), and storage challenges. In addition, the large-scale deployment of hydrogen depends on the availability of low-cost renewable electricity and the development of supporting supply chains.

Furthermore, the high levelised cost of hydrogen production and supply currently makes it uneconomic for process heating applications, particularly at low- to medium-temperature ranges where more efficient alternatives exist (Franco and Rocca, 2024). According to the adapted Hydrogen Ladder version 5.0 (Liebreich, 2025) as shown in Figure 2.1, hydrogen becomes a viable decarbonisation option only in cases where there is a high-temperature requirement combined with process constraints that limit direct electrification (Franco and Rocca, 2024). As a result, hydrogen is generally viewed as a complementary solution rather than a primary pathway for process heat decarbonisation.

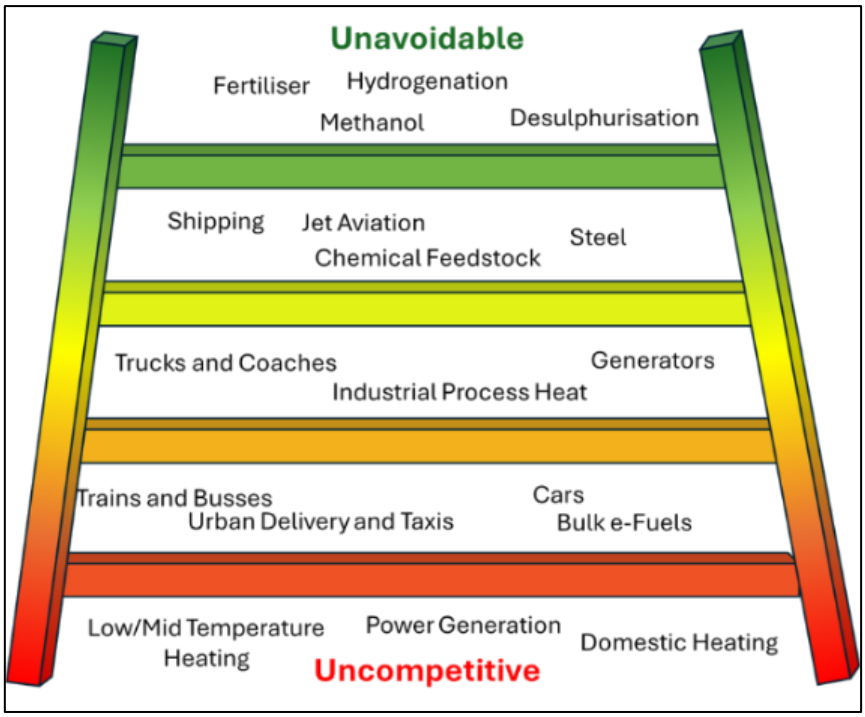


Figure 2.1: Hydrogen Ladder 5.0, adapted by the author based on Liebreich (2025).

In summary, to reduce the carbon footprint of existing direct heating in these sectors, specifically designed heating technologies need to be developed and installed, alongside addressing the major barriers associated with retrofitting alternative pathways into existing processes. For these reasons, high-temperature direct heating is beyond the scope of this macroscale energy model.

2.2.1 Indirect Heating Applications

In contrast, common indirect heating typically uses steam or another medium to transfer heat to the target (Lenz *et al.*, 2020). The heating range of the indirect mode is usually dependent on the medium of heat transfer, while electrode boilers usually supply temperatures of up to 500˚C by providing superheated high-pressure steam (Münnich *et al.*, 2022). Similarly, the biomass boilers also reach 500˚C by combusting biomass material. Another highly efficient technology used for heating is heat pumps. Heat pumps apply the PtH concept and have a coefficient of performance (COP) greater than 1, meaning that they have efficiencies of more than 100% for heating.

In general, most process heat requirements in industries are low to intermediate unless they are energy-intensive operations dealing with minerals, chemicals and metals. Figure 2.2 presents the magnitude of the heating requirement for each industry in New Zealand (EECA, 2024). In New Zealand, total industrial process heat demand was estimated at 115.74 PJ in 2022, of which 68.2 PJ was classified as low and intermediate temperature heat, accounting for approximately 60% of total process heat demand. However, by excluding energy-intensive industries such as metal, petroleum, chemical and non-metal mineral sectors, which are limited to a small number of locations nationwide, the remaining process heat demand is 68.0 PJ. Among the total demand, 21.3 PJ falls within the low-temperature range, while 45.9 PJ is classified as intermediate-temperature heat. High-temperature process heat, defined as temperatures above 300°C, accounts for only 0.81 PJ. This value corresponds to process heat categorised only as above 300°C, with no further disaggregation, which means that the specific temperature level is unknown.

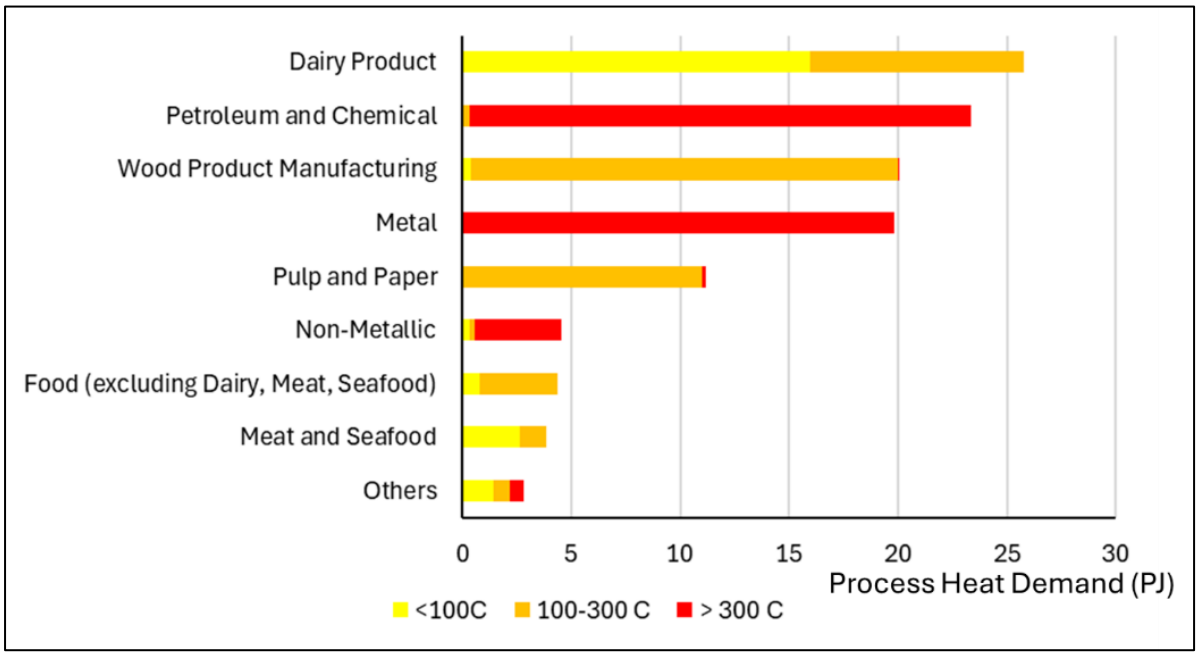


Figure 2.2: Energy End Use Database (EEUD) of New Zealand, compiled by the author based on EECA (2024).

To dive into the temperature levels associated with the processes for each industry, Figure 2.3 shows the temperature ranges of industrial processes under the Statistical Classification of Economic Activities in the European Community (NACE) code (European Commission, 2008). Referring to the figure, most processes have low operating temperatures, except operations which involve drying, evaporating and baking. Indirect heating is a significant field of study as it possesses the properties of how heat was produced since it does not affect the manufacturing process pathway. Therefore, the strategy of fuel-switching and selecting appropriate heating technology is helpful to minimise system cost within the boundary constraints. This characteristic makes it a relatively accessible opportunity for the energy transition.

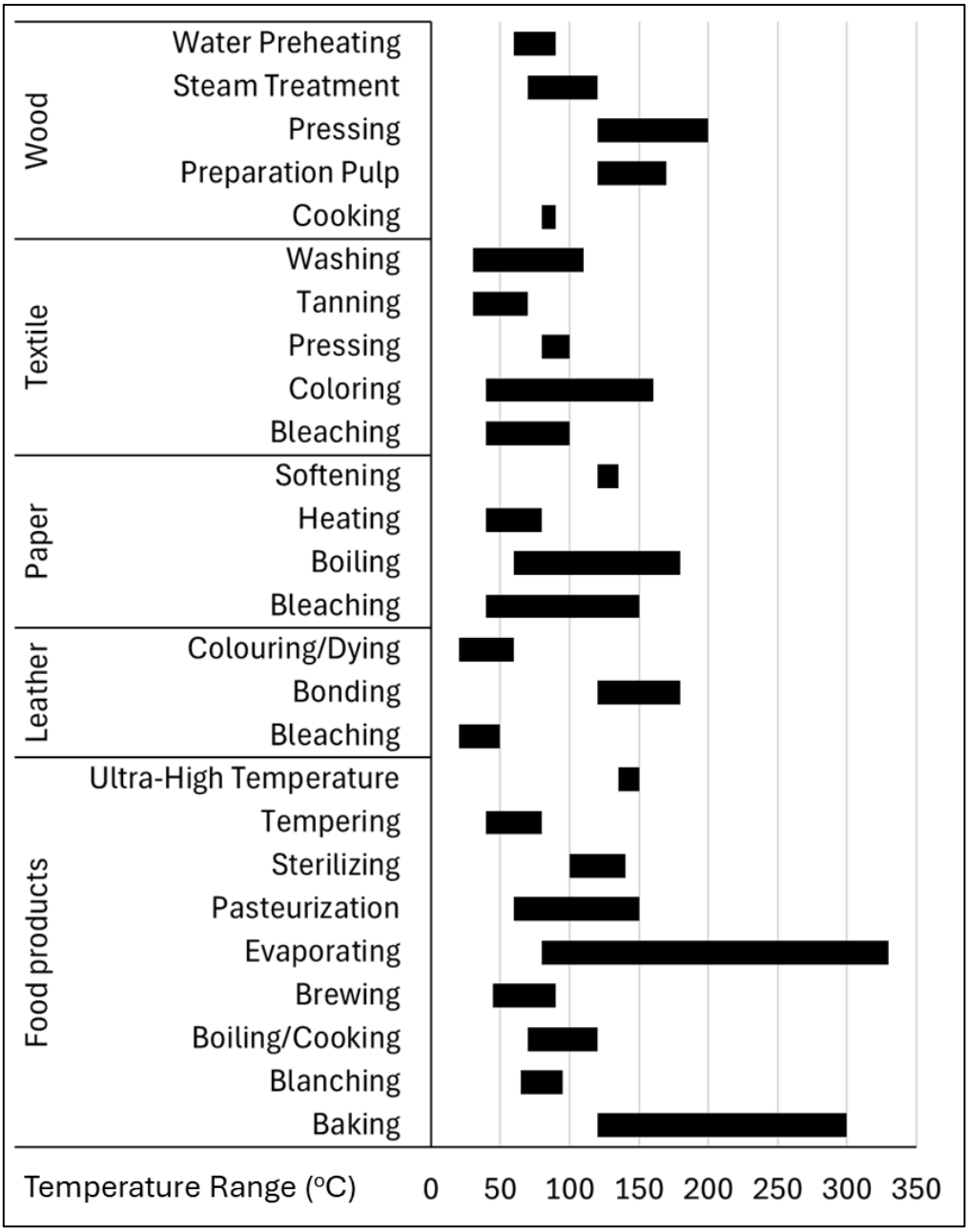


Figure 2.3: Temperature ranges of selected processes based on NACE classifications, prepared by the author using data from the European Commission (2008).

Literature namely von Thadden del Valle et al. (2024) has conducted studies to match temperature requirements with heating technologies for decarbonising process heat in industry. This was done from both a bottom-up and a top-down approach. This section provides a top-down overview of industrial sectors. The approach is important to study the fuel use and heating technology use which are compatible with before conducting any mathematical formulation. As an example, in the pulp and paper mill industry, high-pressure steam is generated to break down lignin in wood chips (Zhao *et al.*, 2019). A bulk heat is supplied by natural gas, supplemented by bioenergy sources including black liquor, which is a by-product from the kraft pulping process (Naqvi, Yan and Dahlquist, 2010). A significant characteristic of the kraft pulping process is the co-generation firing to produce both electricity and heat on-site. Therefore, though direct

electrification for heating in pulp and paper mills is plausible, it is considered more practical and compatible to use bioenergy as a heat source (Lipiäinen *et al.*, 2021). This is because bioenergy is a mature and widely implemented decarbonisation pathway for the pulp and paper industry, supported by the fact that it is readily available in the sector and compatible with the existing process heat system (Joyo, Nastasi and Astiaso Garcia, 2025). A bottom-up approach evaluates the heating requirements and temperature ranges for each unique process. Figure 2.4 shows the heating technologies and their respective operating temperature ranges.

Among the technologies applicable to indirect heating, heat pumps, electric boilers, and biomass boilers are the most relevant options for large-scale process heat decarbonisation.

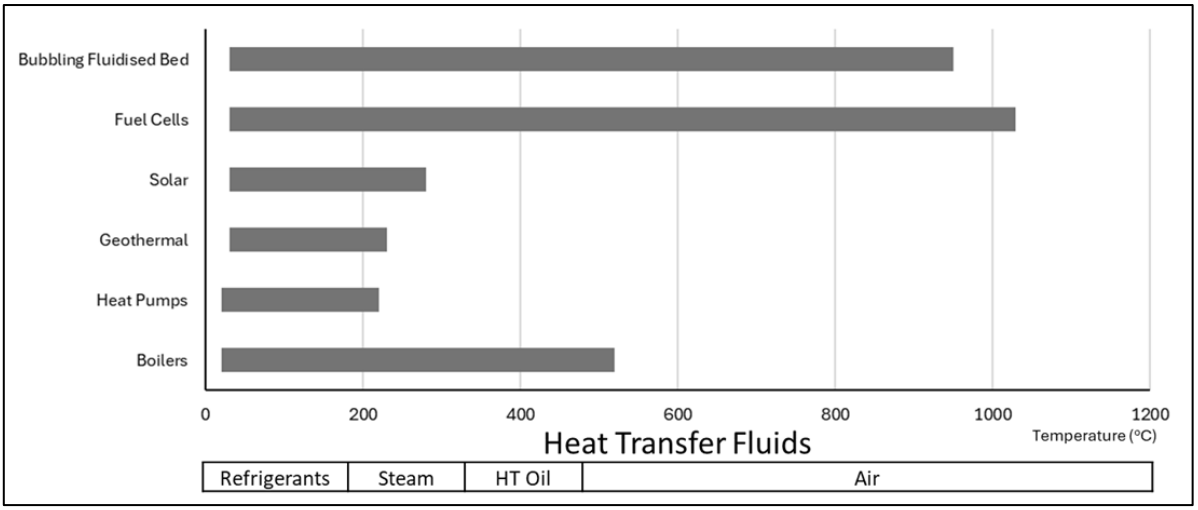


Figure 2.4: Overview of indirect heating technology selection in the market, adapted from von Thadden del Valle et al. (2024).

2.2.2 Processes for Heating Technologies

A heat pump is one of the key technologies for transitioning from non-renewable energy sources to more sustainable alternatives. The primary energy savings achieved through the implementation of a heat pump are affected by several factors, such as (i) the coefficient of performance (COP) attained by the operation, (ii) the current energy source being substituted, and (iii) the fuel source converted to electricity which is powering the heat pump (Marina *et al.*, 2021). Heat pumps are widely adopted for heating in residential and commercial buildings, as they come with large energy savings. However, uptake of heat pump technology in industrial process heat applications has been slow due to the affordability of other heating fuels (Koundinya and Seshadri, 2022).

Figure 2.5 illustrates a system line and the temperature-entropy plot of the thermodynamic operation of an electrically driven single-stage vapor compression heat pump (Bever *et al.*, 2024). A heat pump typically constitutes a compressor, condenser, expansion valve and evaporator. In process 1-2, the working refrigerant vapour is compressed, then in process 2-3 the refrigerant is condensed fully into liquid (Koundinya and Seshadri, 2022). Heat is removed in the condenser for heating. In process 3-4, isenthalpic process occurs by reducing the pressure in the expansion valve.

Heat is absorbed from ambient or waste heat source at the evaporator in process 4-1 as the boiling point of the refrigerant decreases.

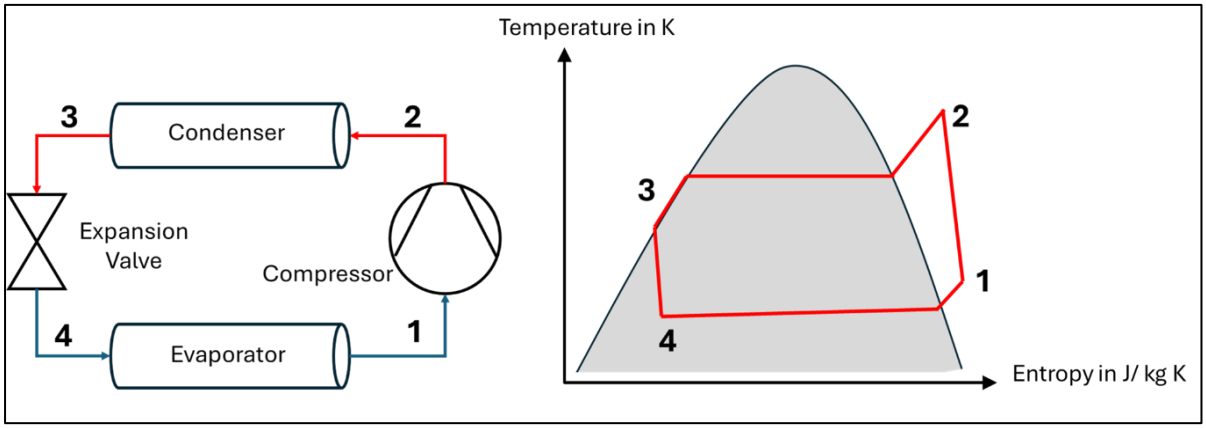


Figure 2.5: Heat pump cycle, adapted from Koundinya and Seshadri (2022).

Equation (2.1) provides the Lorenz COP, which is the greatest theoretical COP for an ideal heat pump running between process and waste heat streams with constant heat capacity and variable temperatures (temperature glides). Equation (2.2) defines the log mean temperatures of a stream. The constant temperature of the waste heat source or process heat sink can be used in place of the log mean temperature for a heat stream without temperature glide (Marina *et al.*, 2021).

$$COP_{Lorenz} = \frac{T_{P,LM}}{T_{P,LM} - T_{W,LM}} \quad (2.1)$$

$$T_{LM} = \frac{T_{in} - T_{out}}{Ln(\frac{T_{in}}{T_{out}})} \quad (2.2)$$

The actual COP is reduced to a fraction of the Lorenz COP value in practical thermodynamic processes due to many losses. Equation (2.3) provides an efficiency term that connects the actual COP to the maximum Lorenz COP (Marina *et al.*, 2021).

$$\eta_{Lorenz} = \frac{COP}{COP_{Lorenz}} \quad (2.3)$$

$$COP = \frac{Q_P}{Q_W} \quad (2.4)$$

The COP is also defined by the ratio of heat output to the electrical power input, as shown in Equation (2.5) (Marina *et al.*, 2021):

$$COP = \frac{Q_P}{E_e} \quad (2.5)$$

With the growing improvement of heat pumps and with heat pumps pushing to medium to high-temperature ranges, the adoption of direct electrification using this technology will accelerate to decarbonise industrial heating. Arpagaus et al. (2018) conducted a market analysis on heat pumps

and deduced that heat pumps are cost-effective up to 100˚C, but the heat sink temperatures above 140^{o} are still in the prototype phase. However, research and development in high-temperature applications is progressing rapidly, as there is strong market potential for heat pumps capable of producing low-pressure steam at temperatures above 120˚C (Bless *et al.*, 2017). In 2023, IEA HPT Annex 58 which is an international collaboration tasked to accelerate the development of high-temperature heat pumps (HTHPs) have advanced the technology readiness level up to 8-9 with heat sink temperatures to more than 120˚C (Zühlsdorf *et al.*, 2023). This shows that heat pumps are a rapidly growing technology as they are seen as a technology to decarbonise industrial process heat. Table 2.3 gives a list of the specifications of high-temperature heat pumps with a technology readiness level (TRL) of 9.

Table 2.3: List of new heat pump technologies and their specifications (Zühlsdorf *et al.*, 2023).

Supplier	**Compressor type**	**Working fluid**	**Capacity (MW)**	**T_{max} supply**
Fuji Electric	Reciprocating	R-245fa	0.03	120˚C
Mayekawa (EcoSirocco)	Reciprocating	R-744	0.1	120˚C
Kobelco Compressors Corp. (SGH165)	Twin-screw	R-245fa/R-134a, R-718	0.4	175˚C
Kobelco Compressors Corp. (MSRC160L)	Twin-screw	R-718	0.8	175˚C
Spilling	Piston	R-718	1.0-15.0	280˚C
Epcon	Centrifugal fans and blowers	R-718 and process vapors	0.2-100.0	210˚C
Everllence	Centrifugal turbo with expander	R-744	10.0-50.0	150˚C
Qpinch	Chemical adsorption heat transformer	R-718, H3PO4 and derivatives	>2.0	230˚C
Ochsner Energie Technik	Semi-hermetic screw	R1233zd	0.5	130˚C

Figure 2.6 illustrates the coefficient of performance (COP) of a range of industrial heat pumps as a function of their respective temperature lift. The Carnot efficiency of most heat pumps lies within the range of 40% to 60%. Heat pumps perform particularly well when the temperature lift (the difference between the sink temperature and the source temperature) is low. When higher temperature heat is supplied using ambient air as the heat source, the temperature lift increases and the COP decreases. Therefore, for indirect heating applications with higher temperature requirements, alternative heating technologies may be more appropriate.

Fossil fuel-based boilers dominated new heating system sales in 2020, with a market share of more than 90% in countries such as Italy and Belgium and total sales of 5 million in the European heating industry (European Heating Industry (EHI), 2021). The heating technology requires a

major transition with different fuel sources for rapid decarbonisation. Among them, electric and biomass-based boilers should be adopted in the strategy plan for higher temperature heating.

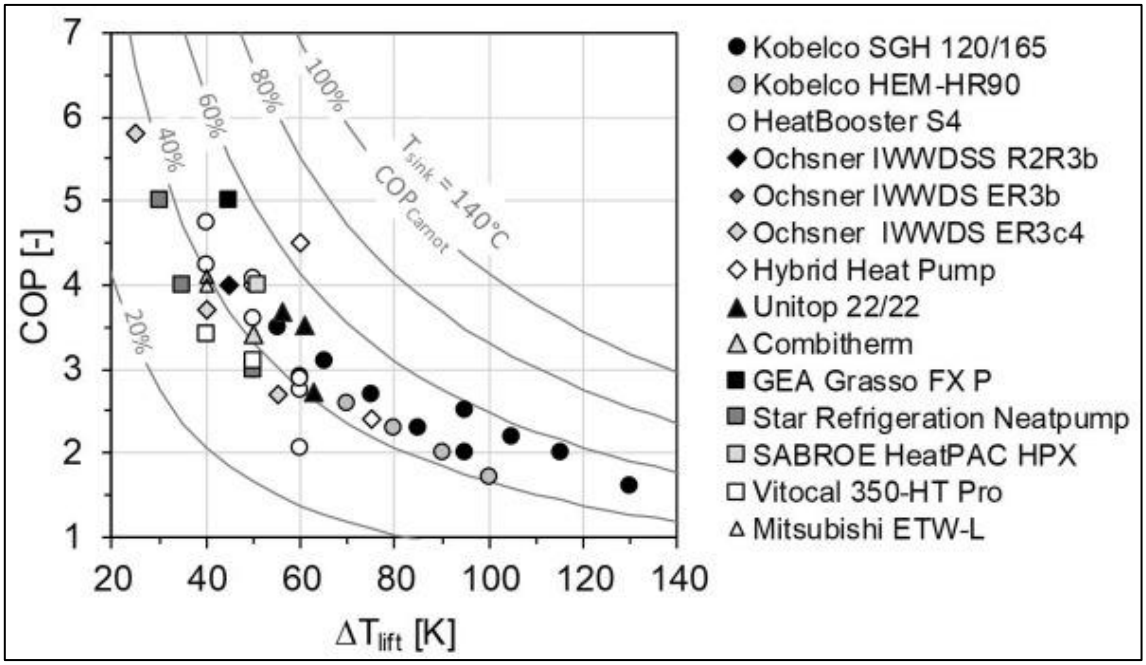


Figure 2.6: Coefficient of performance (COP) of heat pumps as a function of temperature lift, reproduced with permission from Arpagaus et al. (2018).

Electric boilers are a mature and commercialised technology that generate steam or hot water by converting electricity directly into heat (EECA, 2019). Heating is done by passing an electric current through the water in electrode boilers or by using submerged resistive heating elements in electric resistance boilers. Electrode boilers are generally able to achieve high output capacities of up to 98.2 MW with operating efficiencies of close to 100% (Cleaver-Brooks, 2021). Electrode boilers are typically more compact than fossil-based boilers, which allows multiple units to be installed in parallel configuration when substituting a single fossil-based boiler (Schoeneberger *et al.*, 2022). Similar heating effects can be observed for electrode boilers as compared to solid fuel boilers, with the deviation of having water as the heating medium. One advantage of electrode boilers is the faster ramp rates than solid fuel boilers, meaning that they can respond to fluctuations in demand in a more flexible manner (EECA, 2019). The flexibility of electrode boilers is proven by a study made by (Dancker *et al.*, 2019) who optimised the operation of an electrode boiler with thermal storage in response to the volatile energy prices.

Biomass-based boilers are often used by co-firing with coal, which is used commonly for co-generation as part of carbon abatement strategy (Guo, 2022). However, in a net zero carbon scenario, process heat demand should only use sustainable fuel sources. Biomass fuels can be constituted with different kinds of materials, such as in-forest harvest residues, wood briquettes, municipal wood waste, horticultural residues, sawmill chips, agricultural residues, K-grade logs and many more (Hall, 2024). Biomass fuel selection for specific boiler types is primarily determined by calorific value, moisture content, ash composition, and particle size. The selection of boilers can prevent cases such as fuel handling issues, poor efficiency, unscheduled maintenance and high lifecycle cost.

There are two common boiler types that are commercially deployed as industrial heating technologies: bubbling fluidised bed boilers and biomass-fired grate boilers. A high moisture content and high ash content biomass fuel is more suitable for a bubbling fluidised bed boiler or commonly known as a bubbling fluidised bed reactor. This is because a bubbling fluidised bed reactor is flexible with a wide variety of fuels due to the intense mixing of the flow of air through the reactor bed (Gungor, 2013). This characteristic promotes high heat transfer and uniform temperature distribution of the bed. As low calorific value, high ash content and wet biomass fuel are used in a bubbling fluidised bed, the fuel would produce a considerable amount of slagging and deposition after the combustion process (Liu *et al.*, 2020). Hence, the combustion temperature is slightly lower than a conventional coal fired boiler with temperatures at 750˚C and 950˚C to reduce slagging and deposition in the bed (Liu *et al.*, 2020).

Biomass-fired grate boilers are common systems that use a metal grate to hold and combust biomass fuel. Biomass fired grate boilers, or biomass boilers in short, generate heat efficiently. However, they require more uniform and drier fuels, such as wood pellets or wood chips. Therefore, woody biomass is generally the feedstock used in a biomass boiler. For example, regions with large forest cover, such as in the Fukushima Prefecture, Japan, it was assessed that biomass boilers using woody biomass would be competitive for heating after the supply chain is established for co-generation (Schoeneberger *et al.*, 2022).

The design for biomass boilers and coal-fired grate boilers is nearly identical as shown in Figure 2.7, hence there are two options for decision makers for a coal boiler, either retrofit existing coal-fired boilers or install a new biomass boiler. Boiler conversion requires modifications for the fuel feeding system due to different solid fuel types, safety and compliance due to new emission profiles, and combustion chambers due to the change in combustion efficiency. In reality, the fuel-switching for these two options is carried out at the same time. In most instances, it is always cheaper to upgrade the original coal boilers although the depreciation cost of a new biomass boiler is lower (Guo, 2022). Guo (2022) has conducted sensitivity analysis and concluded that retrofitting coal boilers is always more economical than introducing new boilers.

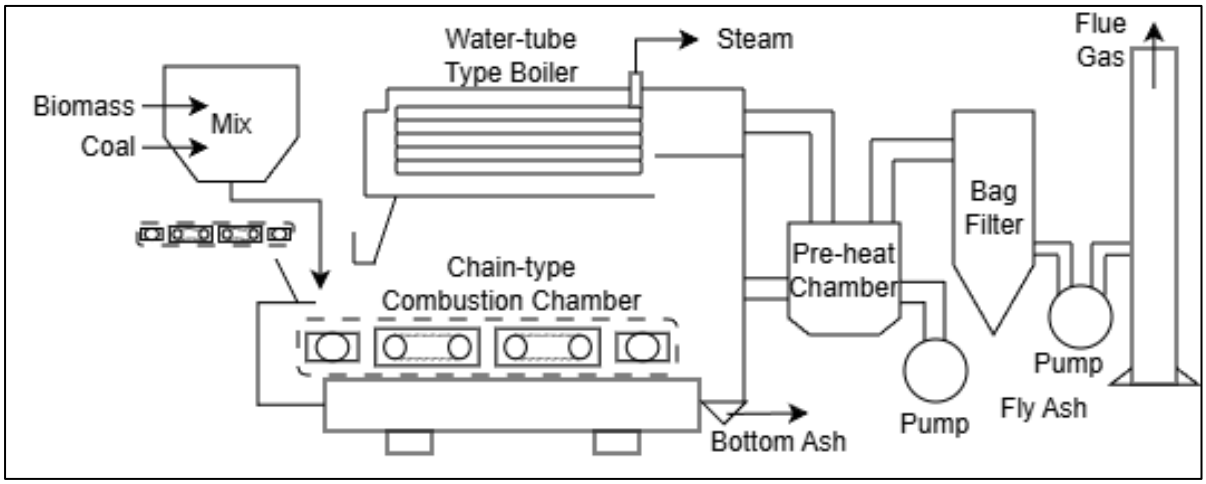


Figure 2.7: Schematic diagram of a coal-fired or biomass-fired grate boiler design, adapted from Huang et al. (2023).

As a summary, the previous section shows that several industrial processes share similar temperature patterns for process heating. This highlights a gap in this section by first possibly categorising temperatures based on generalised processes and subsequently matching them with the heating technology that offers the best rate of return or thermal efficiency.

2.2.3 Heat Demand and Temperature as Modelling Parameters

Temperature is widely recognised as the most important parameter when analysing industrial process heat demand. Apart from heat demand, the required temperature for heating varies substantially across industrial sectors, ranging from metal and mineral production, which can exceed 1000˚C to low-temperature applications in food industries, where the tempering process only requires 40 to 80˚C (von Thadden del Valle *et al.*, 2024).

Studies that quantify temperature levels typically classify heat demand into several categories. However, there is no uniform benchmark for disaggregating heat demand by temperature band, as each study follows its own definitions and classification thresholds. In general, based on the literature studies, a temperature less than 100˚C is usually described as low temperature, whereas temperatures higher than 1000˚C is categorised as the highest temperature band.

Naegler *et al.* (2016) disaggregated industrial heat demand by temperature level using a top-down, sectorial-based approach, while also combining national energy statistics with process-specific temperature shares. The heat demand is then disaggregated into temperature levels by applying industry-specific temperature share profiles from process-level analysis. Hence, the heat is allocated into resolution with four temperature bands: <100˚C, 100˚C to 500˚C, 500˚C to 1000˚C and >1000˚C. This approach is more suited to the purpose of energy system and policy modelling, which is unrelated to site-level decarbonisation.

Bataille *et al.* (2018) explicitly structure industrial demand into three temperature levels but do not quantify them as sectoral heat demand shares. Although applying a top-down approach, temperature levels are used as a decision framework to map decarbonisation technologies to industrial processes. The paper defines three temperature ranges as low, medium and high temperatures. For example, low temperature refers to hot water, space heating and steam generation for heating processes, driven by heat pumps and boilers. Despite providing temperature based on technology compatibility, the paper does not provide any quantitative temperature for heat demand. This omission leaves a gap in the literature for site-level evaluation of fuel-switching options for heating technologies.

Temperature is handled indirectly in the study which focuses on biomass boiler compatibility (Malico *et al.*, 2019). Similar to studies by Bataille *et al.* (2018) and Malico *et al.* (2019) treat temperature implicitly through biomass boiler technology classes and steam applications. However, Malico et al. (2019) show the importance of assessing the technology feasibility and fuel compatibility analysis. Other literature sources, such as the Australian Renewable Energy Agency (ARENA), Madeddu, the IEA, Energy Efficiency and Conservation Authority (EECA), Ronald, Marina, and Mayer, define explicit temperature bands for industrial heat demand. In contrast, studies by Philibert (2017) and Lenz et al. (2020) group temperature levels implicitly using

qualitative classifications rather than quantitative thresholds. Table 2.4 shows the temperature levels of industrial heat across literature.

Table 2.4: Temperature levels of industrial heat across literature sources.

Literature Studies	Temperature classification (℃)						Role of Temperature
Naegler et al. (2016)	<100	100-500		500-1000	>1000		Disaggregate industrial heat demand and assess electrification
Philibert (2017)	<150	150-400	400				Distinguish temperature processes and process transformation requirements
Bataille et al. (2018)	<250		250-1000		>1000		Used as technology-screening framework rather than a demand variable
Malico et al. (2019)	<100	100-200	200-500	500			Biomass boiler technologies and sectoral steam applications
Lovegrove et al. (2019)	<150	150-250	250-800	>800			Technology compatibility
Madeddu et al. (2020)	<100	100-400		400-1000	>1000		A demand attribute to evaluate electrification feasibility
Lenz et al. (2020)	<100	100-200	200-500	500-1000	1000-1500	>1500	Feasibility threshold for biomass applicability
International Energy Agency (IEA), (2018)	<400			>400			Match process heat demand with suitable renewable technologies
Krishnan and Greed, (2023)	<100	100-300		>300			Assess fuel-switching and electrification potential
Muhumuza and Eames, (2022)	<150	150-230	230-650	>650			Potential of waste heat recovery
Papapetrou et al., (2018)	<200		200-500	>500			Optimisation of waste heat recovery.
Marina et al., (2021)	<200		200-500	>500			Match process heat demand with suitable renewable technologies
Mayer et al., (2024)	<200		200-800		>800		Match process heat demand with suitable renewable technologies

As a summary, although many studies explicitly disaggregate industrial heat demand into temperature levels and quantify demand by sector, several reviews and system-level modelling studies treat temperature implicitly as a technology feasibility constraint or omit heat demand quantity entirely. These studies show the significance of having temperature bands with better resolution, especially for low to medium heat pumps with varying temperature lift. These limitations highlight a gap between technology efficiency levels, technology representation, and process-level detail in high-level energy system modelling.

2.3 Green Methanol as a Sensitivity Case

Power-to-fuel pathways have received increasing attention as part of the broader energy transition, particularly where renewable electricity can be converted into transportable energy carriers or chemical feedstocks. Among these pathways, green methanol is of growing interest due to its role as both a fuel and an industrial feedstock. However, in the context of this thesis, green methanol is not treated as a direct process heat decarbonisation technology. Instead, it is introduced as a sensitivity case because it represents a competing demand for renewable electricity.

Green methanol production is typically derived from electrolysed hydrogen, combined with a carbon source for methanol synthesis. As a result, the pathway is electricity-intensive and may place additional pressure on the renewable electricity supply. Under such conditions, green methanol production may compete with process heat fuel-switching for the same renewable electricity resource.

Green methanol is treated as a sensitivity case rather than a core technology pathway in this thesis. Its inclusion allows the model to test how additional electricity demand from an emerging Power-to-Fuel pathway may influence system costs, technology selection, and electricity infrastructure requirements. Framing green methanol in this way helps to assess the robustness of industrial process heat decarbonisation pathways under broader energy transition pressures.

Pratschner et al. (2023) conducted a techno-economic assessment of a green methanol production system. Their analysis includes fixed capital investment for major components and subprocesses, including an alkaline electrolyser, methanol synthesis unit, biomass combined heat and power process, carbon capture, and grid connection infrastructure. The study also incorporates amortised fixed operating costs and variable operating costs to estimate the levelised cost of methanol for a power-to-methanol system. These findings provide useful reference values for cost parameterisation, particularly because mixed-integer linear programming models are sensitive to assumptions on capital and operating costs.

Biomass to methanol pathways

The typical biomass to methanol pathway consists of multiple unit operations, including gasification, syngas cleaning, autothermal reforming, steam electrolysis, methanol synthesis, and methanol purification (Rajaee *et al.*, 2024). Biomass is often dried prior to gasification to reduce moisture content and improve conversion performance. Poluzzi et al. (2022) describe drying

biomass using a belt dryer to reduce moisture content to approximately 15% before feeding it to the gasification reactor. Gasification can be implemented using either direct or indirect configurations. In direct gasification, partial oxidation occurs within the gasifier, typically improving overall efficiency but affecting hydrogen yield. In indirect gasification, heat is supplied externally, which can improve syngas composition for some applications but changes efficiency and yield trade-offs. Poluzzi et al. (2022) report that direct gasification can achieve higher fuel efficiency than indirect gasification, while indirect gasification can yield higher raw syngas flow rates.

Syngas cleaning is required to remove impurities that are corrosive or detrimental to downstream catalysts. Raw syngas is commonly cleaned using a combination of scrubbing and adsorption steps, such as a water scrubber and activated carbon bed, to remove contaminants including sulphur compounds, chlorine, and ammonia (Rajaee *et al.*, 2024). Following cleaning, syngas can be conditioned through reforming. Rajaee et al. (2024) describe catalytic autothermal reforming with an operating temperature around 900°C, which increases hydrogen (H_2) and CO content by converting methane and other hydrocarbons formed during gasification. Literature suggests that autothermal reforming can convert a large share of methane (CH_4) and hydrocarbons, improving syngas suitability for synthesis.

A common method for adjusting syngas composition is the water gas shift (WGS) reaction, which increases hydrogen content at the expense of CO. However, for configurations that can supply additional hydrogen via electrolysis, WGS may be less attractive because hydrogen can be added without sacrificing CO. Butera et al. (2020) reported that solid oxide electrolysis cell (SOEC) systems can achieve higher overall efficiency than WGS-based configurations when hydrogen enrichment is performed via electrolysis. Therefore, the system boundary and hydrogen sourcing method influence whether WGS is desirable. The biomass to methanol plant is presented in Figure 2.8.

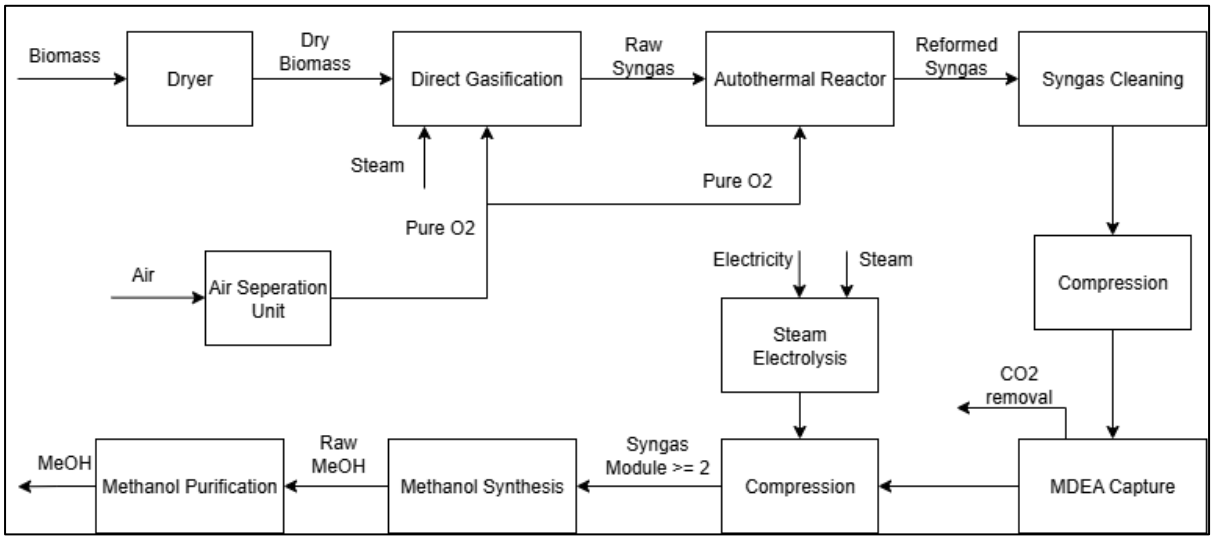


Figure 2.8: Block diagram of a biomass-to-methanol plant, adapted by the author from Poluzzi, Guandalini and Romano (2022).

Methanol synthesis requires the syngas composition to satisfy a stoichiometric constraint, often expressed using the module M, where methanol synthesis is feasible when M exceeds a threshold value (Poluzzi, Guandalini and Romano, 2022). This is commonly expressed as:

$$\mathrm{M} = \frac{(H_2 - CO_2)}{(CO + CO_2)} \qquad (2.6)$$

In short, the green methanol industry is implemented in the energy modelling framework as a sensitivity case because it represents a competing demand for renewable electricity. This would create scenarios in the process-heat-focused energy modelling framework where competing biomass resources would serve as a carbon source for the production of green methanol, while electricity would provide hydrogen molecules from electrolysis. The competition in resources would deliberately affect the configurations of heating technologies employed in each individual factory. The reason is because both green methanol production and process heat decarbonisation require biomass fuel and electricity.

2.4 Energy System Modelling Approaches

"All models are wrong, but some are useful," as famously stated by the British statistician George E. P. Box in 1976. This observation is particularly relevant in the context of energy modelling, where analysing energy demand and energy-efficient pathways across end-use and transformation sectors is inherently complex and time consuming. Projecting future energy demand for industrial sectors based on efficiency and process pathways is even more onerous and uncertain.

The progression of energy models can be split into four different chronological timelines, each characterised by its distinct methodological developments. These periods include the 1970s, 1980s, 1990s and 2010s. The early adoption in the 1950s focused on energy accounting and engineering analysis. In the 1970s, all Organisation for Economic Co-operation and Development (OECD) countries witnessed a substantial economic growth, which led to concerns of rising and uncontrollable energy consumption in those countries (Jochem *et al.*, 2024). This period was also marked by the oil crises, which changed the direction of energy modelling research and energy policies.

Energy efficiency policies backed with technologies emerged in the 1980s, whereas in the 1990s, linearisation of energy markets meant that efficiency improvements were still a priority, but with the addition of a new driver, climate protection (Jochem *et al.*, 2024). With the development of computational power, the 2010s still maintain the legacy drivers while also include sector coupling, electrification and more comprehensive frameworks for analysing the energy transition. Over decades of global events and developments, research directions in the energy modelling space have also changed. As a result, models are increasingly being required to become more comprehensive and structurally detailed to achieve net zero emission targets (Pfenninger, Hawkes and Keirstead, 2014).

One of the earliest optimisation-based models was the Brookhaven Energy System Optimization Model (BESOM) which is the first optimisation-based model developed by Brookhaven National Laboratory (Kydes, 1980). BESOM is developed in the United States to help evaluate energy technologies and policies within the national energy system. It is a basic linear programming model, which provides a natural link between energy security and economic analysis.

The bottom-up models were pioneered in the 1980s with an organised framework which not only accounts for technology choice and adoption, and also the availability of methodological approaches such as accounting frameworks, partial-equilibrium optimisation and simulation models (Fleiter, Worrell and Eichhammer, 2011). Accounting framework-based models are typically distinguished by exogenous definitions of numerous variables. Their goal is to answer hypothetical theories, as they do not include energy costs but is used to evaluate long-term energy demand because of their intuitive and transparency (Chateau and Lapillonne, 1990). Early examples include MEDEE or the revised MED-PRO model (Chateau and Lapillonne, 1990), Long-Range Energy Alternative Planning System (LEAP) (Kuldna, Peterson and Kuhi-Thalfeldt, 2015) and IKARUS (Markewitz *et al.*, 1996).

Leading to this, energy system models which are developed are optimisation-based, meaning that the model is linked to input and output models on a macroeconomic scale. One typical example of an optimisation model is the (MARKet ALlocation) MARKAL modelling framework, which follows a bottom-up approach, and was developed by the IEA's Energy Technology System Analysis Programme (Sweeney, 1981). This optimisation includes an objective function that aims to minimise the total system cost. It also ensures a balance between energy supply and demand, allowing for a clearer understanding of how both interact (Energy Information Administration, 2025).

The evolution of energy modelling has also been strongly shaped by changes in electricity systems. The introduction of non-dispatchable sources in the generation mix meant that demand-side management alone was no longer sufficient, and electricity demand increasingly had to realign with fluctuating wind and solar generation (Jochem *et al.*, 2024). This led to greater focus on system integration and reliability. Therefore, energy models in that period began to incorporate time preference, market heterogeneity, and factors that synthesise various elements of non-rational investment decisions.

There is a growing interest in the literature regarding sector coupling and integrated energy. Research questions were raised on how to effectively integrate primary energies which are intermittent in nature with users such as residential, transport, industry and commercial players by changing the boundary conditions of the users. As an example, substitution of fossil fuels in the energy consumers by electrification, for electrode boilers and heat pumps for process heat. Direct or indirect electrification by short-to-medium term storage and small-scale rooftop photovoltaic panels. Although the terms "sector coupling" and "energy integration" can be ambiguous, they are often used to describe the replacement of fossil fuels with electricity within the Power-to-X concept (Jochem *et al.*, 2024):

- Power-to-Heat (PtH), e.g., heat pumps, electric furnaces

- Power-to-Move (PtM), e.g., flywheel, electricity-powered vehicles
- Power-to-Gas (PtG), e.g., green hydrogen
- Power-to-Liquid (PtL), e.g., green ammonia, green methanol or e-fuels.

Among other things, these developments require high spatial and temporal granularity. Contemporary models which have moved from aggregated regions into partitioned locations, should consider the structural changes in energy generation and demand centres across regions. The resolution can even be improved by having several local centres which represent high energy density nodes (Nägeli *et al.*, 2020). While leveraging the development of highly computational resources, energy models moved from typical weeks or months towards hourly resolutions in a yearly time horizon to capture electricity generation with the spot electricity price (Sensfuß, Ragwitz and Genoese, 2008; Nägeli *et al.*, 2020).

As of 2025, the focus of energy system models is mainly answering the question of energy trilemma, which is security of supply, environmental concerns and sustainability as well as the energy equity (affordability). However, there is a high possibility that the modelling approaches will diversify to address targeted research and policy needs in the future.

2.4.1 Core Dimensions of Energy Models

Energy system models comprise multiple interacting dimensions, including energy commodities, system components, transformation pathways, and spatial and temporal resolutions (DeCarolis *et al.*, 2017). In some cases, stochastic scenarios are also incorporated as an additional dimension, particularly in bottom-up modelling approaches that typically rely on a single optimisation objective function.

Energy systems with sector coupling specifications typically consist of a variety of commodities which can be in the form of energy, energy carriers or even feedstocks, all within the scope of life-cycle assessments (LCA) (Aryanpur *et al.*, 2021). Commodities are usually driven by the supply-side motivation where location from spatial disaggregation provides detailed meteorological information at wind speed and solar irradiation, which are beneficial to analyse the capacity factor and electricity generation from these site-dependent locations (Egerer, 2016). Hence, electricity is an example of a commodity in the power systems model.

Commodities also include solid sustainable fuels such as bioenergy and other utilities such as water and green hydrogen carriers. Availability of biomass and its types in specific forestry areas are important for a hybrid electricity and biomass process heating in industrial plants (Taylor, Atkins and Walmsley, 2025). On the other hand, hydrogen is an energy carrier which can be stored, combusted or to be used as a chemical feedstock to manufacture hydrogen-based compounds. In short, commodities represent the entities that flow between system components that either consume or produce them.

Space is a dimension often covered by many models which address infrastructure issues in a spatial manner. The spatial attributes are typically represented by region, location or nodal level. Single-nodal systems only have a spatial scope of a single location, often referred to as the

"copper-plate assumption", meaning that commodity flows between components without any technical limitations. Single nodal systems usually aggregate specifications and data, often sacrificing details such as capacity-related constraints. Multi-nodal systems typically operate at a higher spatial resolution, which allows capacity limitations in commodity delivery to be represented as constraints within the energy system model. For example, the capacity of a transmission line in the electricity network can be a limiting factor for electricity delivery in a multi-nodal system, whereas such boundary constraints would not be captured in a single nodal system.

The time dimension is not a standalone component within an energy systems modelling framework. It provides temporal detail on the behaviour of system components by representing changes in commodity flows using discrete time steps. Incorporating the time dimension allows fluctuations in demand and the intermittency of energy technologies to be captured, leading to a better understanding of how system behaviour evolves over time. Additional features such as storage can also be included, as they help to stabilise interactions between transient energy supply and demand. Overall, the time dimension enables energy models to assess operational reliability when responding to uncertainty and transformation processes within the system, while also supporting evaluations of energy security under extreme conditions such as dry years or peak demand scenarios.

Components constitute an important dimension in energy system modelling, particularly for representing energy technologies. Each component is characterised by attributes such as capacity and operational capability within the wider energy network. In capacity expansion models, component capacities are typically treated as decision variables which will allow the optimisation to determine appropriate sizing that minimises annualised cost while satisfying operational limitations. Installed capacity acts as a limiting factor on energy flows by linking capacity decisions directly to the system's operation.

Components are also responsible for energy conversion as commodities undergo transformation within the system. Each component comprises a set of capacity and operational variables, often associated with coefficients that represent efficiencies, capacity factors, utilisation factors, or unit conversions. In a bottom-up framework, it is essential to represent energy components explicitly so that commodities can flow from supply centres to demand centres. The resulting energy flow network can be effectively represented as a graph, where nodes correspond to spatial locations or technology components and connections represent energy flows. In general, component types in an energy system model include sources and sinks, converters, electricity lines, storage units, generators, and heat exchangers.

Transformation Pathway

Energy system models often cover a defined planning horizon, typically ranging from ten to ninety years (Aryanpur *et al.*, 2021). Transformation pathways are used to explore optimal transition routes as long-term strategies, with the aim of informing decision-makers on energy planning and policy implementation over the studied period. This dimension is important for assessing changes

in the energy landscape through a stepwise transformation approach, which reflects the practical realities of periodic transition initiatives.

With the inclusion of transformation pathways, the lifetimes of system components can also be considered. This process includes the decommissioning of existing technologies, which reduces the total available capacity over time and results in a transient capacity stock within the system (Hoffmann *et al.*, 2024).

Most components are directly linked to total system costs, as increases in capacity sizing are often proportional to technology costs. In some cases, a fixed cost is also imposed when a unit operation is selected as a discrete decision. As a result, components play a major role in shaping transition pathways with their underlying cost implications and energy conversion characteristics. Hubs within a spatial region are used to ensure that commodity flows are balanced within that particular region, but they do not usually represent network components that drive system costs. As a result, each aggregated spatial region is assumed as a single copper plate with no internal constraints.

Although space and components share some basic characteristics, such as being represented by nodes in a graph and influencing capacity variables and operational outputs, they serve different roles in the energy system model. Components are consistently defined as distinct classes, each equipped with specific attributes, functionality, and a defined purpose within the model. Accordingly, Figure 2.9 illustrates how predefined components operate within a multinodal energy system and demonstrates the internal architecture of a specialised component through class inheritance and associated attributes within the ETHOS.FINE (Groß *et al.*, 2023).

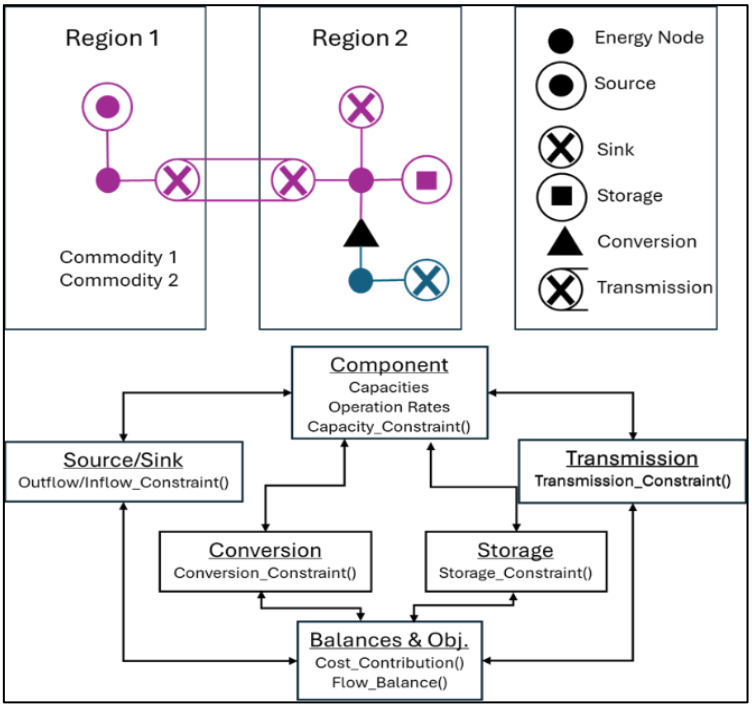


Figure 2.9: A depiction of an energy system model in a directed graph approach and a typical backend architecture on components representing specific respective energy technologies adapted from Hoffmann et al. (2024).

2.4.2 Complexity, Spatial Resolution and Temporal Resolution

As direct experimentation and validation on energy systems is rather impractical, energy system models are employed to study system behaviour (Ridha, Nolting and Praktiknjo, 2020). A wide range of model types and classifications have been proposed in the literature sources (Lopion *et al.*, 2018). It is generally accepted that any energy model represents only a selected portion of the energy system and cannot fully replicate all real-world behaviour. Instead, models capture the properties that are most relevant to the modeller's objectives. As noted by a famous Austrian-born sociologist, (Neurath, 1973):

"To remember simplified pictures is better than to forget accurate figures."

This statement reinforces the idea that simplifying complex systems is often necessary to improve understanding of the behaviour represented by a model. It is important that the concept of complexity is clearly defined within the context of an energy system, as many critical decisions are made based on the energy model. Understanding complexity helps to distinguish whether it arises from the underlying energy system itself or from the way the system is represented in the model (Bale, Varga and Foxon, 2015). In this section, complexity in the model dimensions, including spatial and temporal, are reviewed for simplification pathways. In this section, complexity arising from spatial, temporal, and other dimensions of the model is reviewed in relation to literature to identify potential pathways for simplification.

Given the sheer amount of factors and specifications which have to be considered in the energy model, three basic principles are summarised to handle the complexity in the present published literature (Kotzur *et al.*, 2021).

- Reduction: simplifying the system by decreasing the number of components, variations, and processes, thereby reducing the number of variables and interdependencies.
- Control: managing unavoidable complications effectively through suitable strategies, which includes anticipation of system dynamics.
- Avoidance: preventing unnecessary growth in complexity by eliminating specifications which are not required for the intended analysis.

The scope, spatial resolution and temporal granularity of the energy system model determine the size of the model. The size of the model matrix is first affected by the number of variables, which inherently increases the number of equations which link the variables with the constraints, all integrated within the objective function. The model size usually increases linearly with every number of variables present in the formulation, regardless of the type of variables (continuous, integer or binary) and type of constraints (linear or non-linear) (Pfenninger, 2017).

Spatial Resolution

Spatial extent refers to the geographical scope conducted by a studied scope and serves as a demarcation of territory where the analysis is conducted (Frew and Jacobson, 2016). Depending on the objective of the study, it plays an important role in shaping the structural design of an energy system model. Based on literature sources, the geographical extent can be categorised

into global, regional, national, local and project-specific. Global energy system models typically focus on issues such as anthropogenic carbon concentration (Messner and Schrattenholzer, 2000), emission reduction targets, global mitigation policies and climate impacts on the energy supply chain (Labriet *et al.*, 2015) and implementation of a share of renewable generation (Akashi *et al.*, 2012). Regional models usually involve countries or aggregated countries at a continent level. It is crucial that the term used for region is clearly defined, as it is used to describe two distinct terms in literature. Region could refer to a group of countries or denote subnational units within a country such as a state, territorial authority or counties.

Examples of these include renewable electricity deployment in Sub-Saharan Africa (Barasa *et al.*, 2018), rapid decarbonisation pathways for Europe (Siskos *et al.*, 2018), and energy security in the Baltic region (Augutis *et al.*, 2020). The national-level model includes the supply and demand within a country. The national-level model includes analysis of a certain well-defined sector to study or the energy-related economic activity of the country. Examples include ambitious net zero emission pathways in Japan (Fujimori *et al.*, 2019) and Ireland (Yue *et al.*, 2018), transport systems in Denmark (Tattini *et al.*, 2018) and power generation development in Australia (Rosa and Castro, 2020). A local energy system model further reduces the extent of the scope to a smaller spatial boundary, which would address sub-national areas within a country, such as transport electrification in New York City (Isik, Dodder and Kaplan, 2021). At the most detailed scale, project-level models focus on a single site or a complex, such as the assessment of utility-scale photovoltaic systems (Obeng *et al.*, 2019) or wind farms for electricity generation (Touafio *et al.*, 2020).

The higher the detail in the spatial resolution, the more considerations and links have to be made between nodes, leading to an increase in the number of constraints and variables in a model. It is often dependent on the modellers' choice to incorporate systems design as a consideration factor, as certain constraints have to be introduced which might change the solution structure. This process includes site-specific characteristics and connections such as logistics networks, transmission and distribution lines, and electrical assets within the grid. It also accounts for variations in supply and demand, including differences in generation output and energy use that arise from local conditions, such as climate-dependent wind and solar generation, geothermal availability, and heating or electricity demand influenced by colder climates (Aryanpur *et al.*, 2021).

Reducing the number of model regions would generally lead to fewer grid nodes and a smaller set of available technologies. As a result, it is useful to distinguish between spatial aggregation, which relates to grid topology, and technological aggregation, which concerns the representation of supply and demand technologies.

Spatial aggregation can be defined by a relative locality (Bogdanov and Solomon, 2023) or by coordinates (Phillips *et al.*, 2023). In addition, clustering techniques such as k-means or max-p are often used for spatial aggregation. Several research projects have presented their work on the impact of spatial aggregation of grid constraints within the energy system model (Hörsch and Brown, 2017). Other reports of spatial aggregation include the derivation of a clustering-

equivalent network using k-means by assessing the geometric distance of demand and generation in the European transmission system (Hörsch and Brown, 2017). Literature also includes clustering based on attributes of a technological component in the system, such as market price categorised by zones (Martínez-Gordón *et al.*, 2021), electrical distances, or capacity factors (Frysztacki, Recht and Brown, 2022).

Cao *et al.* (2019) and Frysztacki, Recht and Brown (2022) highlight that clustering methods based on nodal loads or capacities, marginal costs, nodal prices, electrical distances, or radial equivalents are inherently tied to the existing system configurations. Due to the attributes being clustered in the model for simplification, these approaches might present a difficulty in capacity expansion models, as the suggested upgrades heavily rely on the current network structures. Therefore, it is often acceptable for a clustering method to target the simplification of dispatch models in the existing systems. Hence, clustering methods based on capacity factor time series and other size-dependent parameters could be applied to the existing energy system for expansion modelling.

Temporal Analysis

Temporal resolution similarly affects model size and behaviour. The cardinality for temporal resolution could range from decades to months, days, hours or up to minutes (Pfenninger, 2017). Spatial resolution is important to access the structural optimisation of the optimal solution, whereas temporal granularity helps to study the dynamic operational behaviour which would inherently affect the investment decisions of the model. Similar to the spatial and technological resolution, where both would have major implications to the problem size due to the change in the details in each location, temporal resolution would increase the number of constraints and variables within each component by an addition to the dimension.

The number of operational constraints and variables declared in the problem is almost linearly proportional to the number of time steps applied in the model. Each technology or component in the model must have variables and constraints defined for each time step, which are typically linked to the equation governing the capacity shared across all time periods so as to account for investment costs (Hoffmann *et al.*, 2020).

The energy system's complexity stretches over time, temporal resolution, and the time horizon, as mentioned above. The time horizon should be defined to capture the time scale and the transformation pathways and learning rates, whereas the temporal resolution is important for demand response and intermittent non-dispatchable resources (Neniškis and Galinis, 2018).

Generic energy models often use hourly resolution for a yearly horizon in their model, primarily due to the limitations of available datasets from the database possess that time interval, especially when dealing with electricity prices, capacity factor, or electricity demand. However, this level of hourly resolution with 8760 time steps is often unnecessary and inefficient, as it substantially increases problem size without improving data consistency or solver performance (Hoffmann *et al.*, 2024). Although it is often acceptable for smaller-sized models, this approach significantly increases the computational effort for solving. Hence, it is often important to pre-

process the data regarding the time dimension by temporal aggregation before inputting it to the model, as the data is often exogenous.

The concept of temporal aggregation is used to replace a large number of time steps with a smaller number of time steps (Hoffmann *et al.*, 2020). The approach can be implemented by reducing the number of time steps within the model, merging those time steps, and representing them with an average value, or by applying statistical analysis to represent the time steps with a set of generalised periods, as illustrated in Figure 2.10.

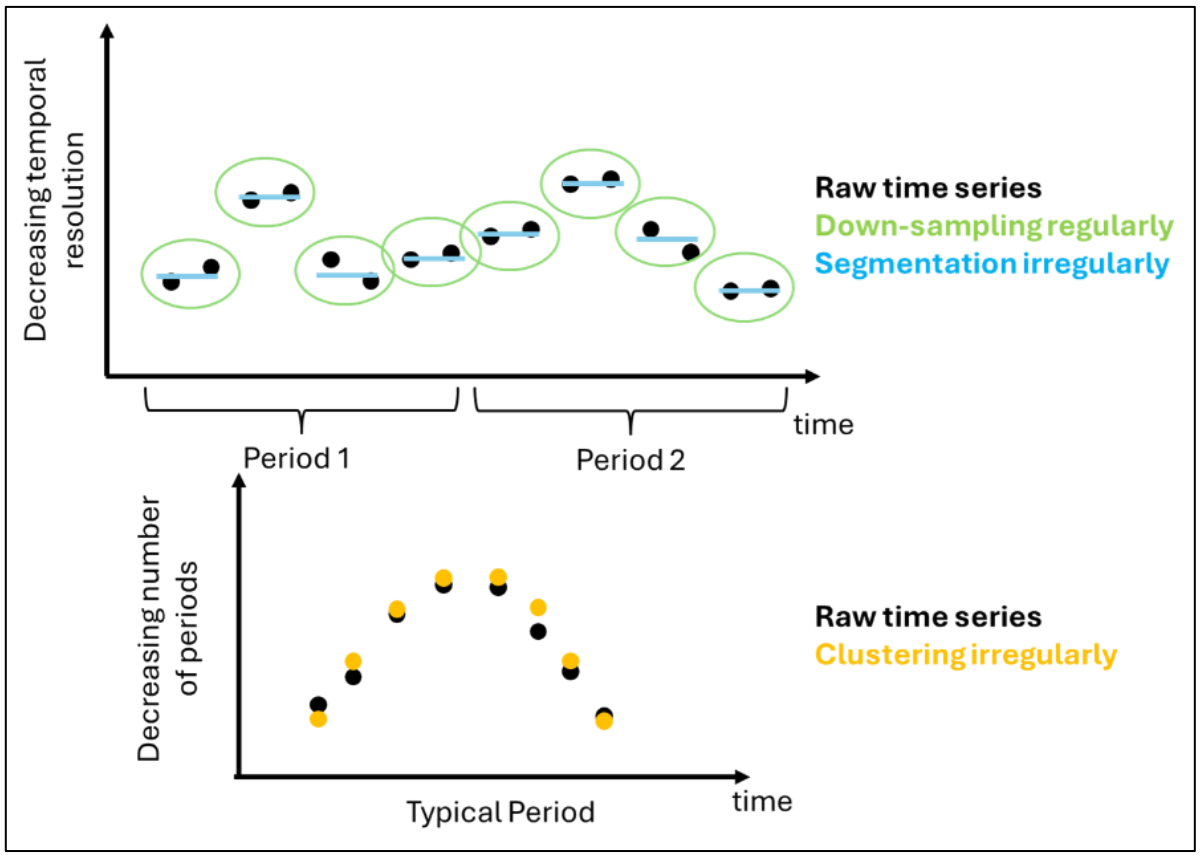


Figure 2.10: Overview of methods for reducing temporal complexity in discretised time space, adapted from Kotzur et al. (2021).

The common method to decrease the total number of time steps in the model is to reduce the temporal resolution altogether. The simplification could be done by segmentation or down-sampling while maintaining the attribute of taking time steps continuously. Down-sampling refers to taking the average or sum of several time steps. As a result, the simplification method works, but it would not reflect the reality of the energy systems model. This issue is often exacerbated by raw input with a large variance, often with inconsistent data affected by uncontrollable uncertainties such as weather, human intervention and many more. Under such applications, required capacity investments tend to be underestimated, while self-consumption levels within energy systems are overstated (Pfenninger, 2017). This process includes inflated estimates of electricity fed in from photovoltaic systems, battery storage, and wind turbines (Beck *et al.*, 2016).

Another approach for reducing temporal is segmentation, which involves combining neighbouring time steps based on their similarity in values. Although the time steps remained continuous, the

method produces time intervals of varying length, as shown in Figure 2.10. Several techniques have been proposed in the literature. Some works rely on a clustering approach, such as partitional clustering or agglomerative clustering, with the restriction that only consecutive time steps may be categorised (Hoffmann *et al.*, 2020). At the same time, other approaches formulate the segmentation problem as a mixed-integer linear program that merges adjacent time steps while minimising deviations from the original time series (Stein *et al.*, 2017).

This segmentation method could also be useful to aggregate smaller time steps into large time steps like hours to days, weeks or even months, such as energy demand, fuel requirements or capacity factors of generations which have a consistent and predictable cycle. This temporal aggregation approach is widely used in many energy modelling frameworks, such as Long-range Energy Alternatives Planning system (LEAP) (Gebremeskel, Ahlgren and Bekele, 2023), Open Source Energy Modelling System (OSeMOSYS) (Slimani *et al.*, 2024), TIMES (Balyk *et al.*, 2022) and many more.

Another method used is the clustering of mutual similarity of typical periods. Simplification using typical periods groups not only adjacent time steps but also similar time segments across the time series. As illustrated in Figure 2.10, time steps that exhibit similar patterns are aligned based on their profiles. The time series is first normalised and then organised into a dissimilarity matrix, from which representative typical periods are identified. Each time period within this matrix is characterised by a multidimensional feature space that is used for clustering. There is a rising interest in the application of temporal aggregation techniques for energy modelling (Bahl *et al.*, 2018). Several studies have been conducted to tackle the temporal aggregation of time series on storage technologies by clustering methods like k-mean centroids (Kotzur *et al.*, 2021) and Pareto fronts (Gabrielli *et al.*, 2019) which are cumbersome to begin with.

Technological Aggregation

Technological aggregation generally leaves the network topology unchanged and therefore introduces no additional connectivity or boundary constraints. The most basic method aggregates all generation sites of the same technology into a single representative site in each region (Phillips *et al.*, 2023). More advanced approaches include applying clustering techniques based on techno-economic characteristics, such as levelised cost of electricity, capacity factor time series, or combinations of site-specific, technology-related, and price-related attributes. A summary of these aggregation methods is provided in Table 2.5.

Table 2.5: Summary of technological aggregation from literature sources.

Authors	Year	Description
Klemm, Wiese and Vennemann	(2023)	Weighted average for PV panels on roof for electricity and thermal uses.
Phillips *et al.*	(2023)	Clustering of sites and setting capacity limit specifications on technological representation on intermittent renewables.
Pöstges and Weber	(2023)	Aggregating wind generation sites and technologies based on investment choices.
Frysztacki et al.	(2021)	Aggregating generation technologies within subregion groups.

A commonly used approach to reduce complexity is called myopic foresight. The myopic foresight approach optimises one investment period at a time. The system design selected for the current period becomes a starting point for the next period. The model performs brownfield analyses, where each optimisation builds upon existing infrastructure rather than starting from scratch (Kotzur *et al.*, 2021). Instead of optimising the system design in a forward time step, another extension called backcasting is created, the model first solves for a desired future target system. The first target is then used to define constraints for the previous investment period. The process is iterated backwards in time until the initial period is reached. Although this method requires one additional optimisation compared to the forward myopic approach, it helps avoid investment decisions that are delayed until later periods, which could otherwise become unnecessarily expensive. Another approach is known as the rolling-horizon approach, where a single optimisation spans at least two periods.

However, the perfect foresight models which consider the entire investment horizon all at once are invariably considered. Several studies have examined the performance of foresight modelling and unanimously concluded that it has a better edge over myopic and rolling-horizon approaches. Using Germany as a case study, Lambert et al. (2023) investigated how different forecast horizons influence the outcomes of energy system simulations. The study shows that shorter forecast horizons, often associated with myopic modelling approaches, can lead to suboptimal investment decisions, such as delayed infrastructure deployment and higher long-term system costs. Longer forecast horizons, on the other hand, give a better idea of what the system will need in the future, which leads to more stable and cost-effective transition pathways. Thomsen et al. (2021) also revealed that the myopic approach cannot compute long-term benefits with large initial investments due to limited foresight.

However, solutions from perfect foresight are less realistic and yield less robust cost projections for technological transformations, as the optimal solution returned is the overall structural optimised solution, leaving out other possibilities of pathways to be present in the model (Löffler *et al.*, 2019). Nevertheless, the perfect foresight approach is adopted despite being more computationally expensive and complex for larger-sized problems. Instead of applying a transformation coupling pathway to reduce model size, model complexity can be simplified by linearisation reformulation.

As the introduction of a unit of technology is related to the price and capacity or performance, it is appropriate to be coupled with a binary or integer variable (Bahl *et al.*, 2018). This approach is computationally intensive because it aims to represent real-world system behaviour in detail. Therefore, depending on the complexity of the problem, the cost function could be non-linear, which is often a general representation of the behaviour of cost per unit decreasing as technology size increases due to economics of scale. Conversely, a convex quadratic cost term is applied for higher investment costs for higher capacities (Lopion *et al.*, 2019). A good example is the deployment of wind turbines, where land availability constitutes a significant portion of the cost; this factor makes deployment economical within limited space but increasingly cost-intensive as further expansion is required, which can be represented using a single quadratic cost function. Referring to Figure 2.11, similar to operational constraints which affect an energy model, the

mathematical program type can be modified depending on the model and scale of the technology. A simplification in the model formulation is common to approximate technology investment by programming the model into a Mixed-Integer Linear Programming (MILP) (Mashayekh *et al.*, 2017). With the existence of a MILP, a binary which shows the existence of the technology, a fixed cost and a proportional cost to the technology can be applied to size a unit operation. Furthermore, a piecewise model could be used, if necessary, to dissect a cost curve into several linear costs to account for different capacity ranges. This approach would lead to a larger number of binary variables but with a good amount of accuracy.

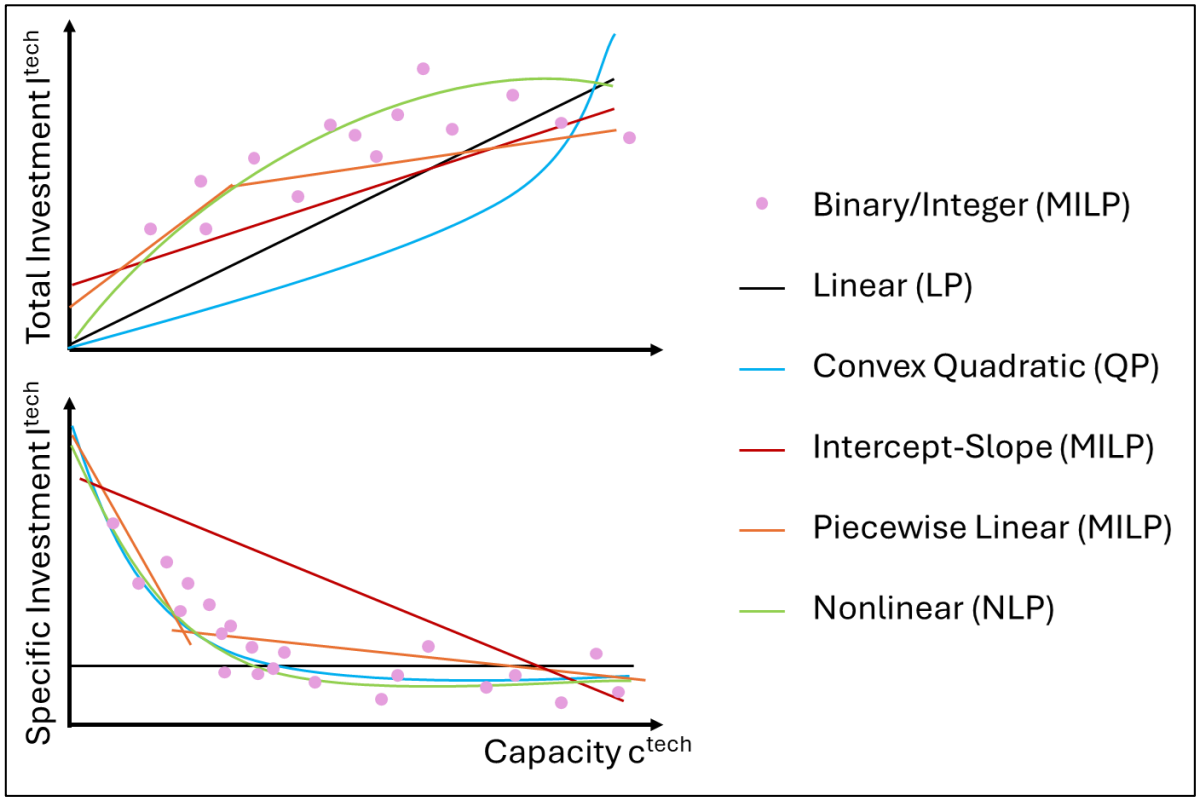


Figure 2.11: Mathematical program variations based on investment cost formulations, adapted from Kotzur et al. (2021).

2.5 Existing Energy Systems Modelling

This review draws on databases for energy-related system modelling tools from the "Open Energy Platform" (Open Energy Platform, 2025) and the "Open Energy Modelling Initiative" (Openmod Initiative, 2025) with 30 and 100 entries respectively for types of tools. Out of the combined 130 entries, 14 tools were overlapped or an extension of the same tools, meaning that there are 116 unique tools available in the database. Out of the tools, more than a hundred are formulated based on linear or mixed-integer linear programming methods.

There has been a strong shift towards open-source energy modelling tools ever since 2010. This trend is largely expedited by public authorities and research funding initiatives such as Horizon Europe, which place priority in making tools transparent with their code base and open access (Jochem *et al.*, 2024). This is due to the initial phase of the energy tools, which were

commercialised and proprietary protected. These energy tools were often designed to be black-box or closed, which makes it difficult to interpret results and draw conclusions on the interactions between technologies in the model. Therefore, most energy models recently being developed and used are open source in response to the drawback from proprietary models (Hörsch *et al.*, 2018).

Available energy modelling tools cover a wide range of applications and functions, including transport planning, district heating, electricity market pricing, energy policy analysis, emission target assessment, energy system transformation, and grid analysis. Their primary purpose is to support the provision of energy services to end users by representing interactions between energy supply and demand, while also accounting for social and environmental considerations within the modelling framework.

Further review of the literature is required to ensure that energy system tools are appropriately defined and classified for energy system analysis, particularly in distinguishing between simulation-based and optimisation-based tools (Connolly *et al.*, 2010). Additional screening is undertaken following the approach of Groissböck (2019), who filters energy system tools according to their capabilities in modelling heating systems and performing optimisation. This filtering process aligns more closely with the objectives of this thesis, which focus on the decarbonisation of industrial process heat and electricity grid modelling at a national scale. In addition, Ringkjøb, Haugan and Solbrekke (2018) provide insights into effective categorisation of energy system tools based on their purpose, scope, and modelling approach, supporting the systematic selection of tools that are suitable for addressing the research questions.

Lahijani, Protheroe and Gschwendtner (2025) have proposed a set of criteria specifications to apply to the list of tools in the selection processes. These bases were created so that the tools selected would have the capability to effectively model industrial process heat systems. In order for a tool to meet a basis, it is important that the tools are documented explicitly, which specifies their claim on demonstrating the task, which is shown in Table 2.6 (Lahijani, Protheroe and Gschwendtner, 2025).

Popular energy models in the market, such as TIMES, MARKAL and OSeMOSYS, are often employed by many literature sources. TIMES provides a detailed technological representation for each component. However, it is mainly focused on system-wide analyses rather than the operational design details of individual industrial process heat systems (Kannan and Turton, 2012). MARKAL plays a strong role in financial optimisation and energy market interactions, making it a valuable tool for backing energy policies. One of the main shortcomings of MARKAL is that it does not deal well with spatial dimensions, meaning that the technologies specified would be a generalised component representing an aggregate system model (Krzemień, 2013). Similar to MARKAL and TIMES tools, OSeMOSYS is designed to study the energy scenarios with the intention of making nationwide or regional energy policies; consequently, the design for site-level industrial or electricity generation nodes is rather impossible due to the lack of granularity (Howells *et al.*, 2011). Overall, these tools are optimised for strategic planning and aggregated studies, making them useful for analysing broad trends and long-term projections. They provide valuable insights

before decisions on large-scale energy system design are undertaken. Therefore, these tools do not pass the criteria of location-specific and capacity-sizing criteria, which are indispensable for a site-level decarbonisation effort.

Owing to the breadth of existing literature on energy modelling tools, a comprehensive review is beyond the scope of this thesis. Only a representative selection of tools with direct relevance to the research objectives is considered. This selection for evaluation is made only for the tools which fulfil the basis which was defined in Table 2.6. The shortlisted energy tools are OEMOF, TRNSYS, EnergyPLAN and Calliope.

A good energy model must consider several features. For the input section, demand profile generation, data input flexibility and embedded functions and libraries are qualities of a satisfactory energy model. For the case of OEMOF (Hilpert *et al.*, 2018), it offers a high degree of flexibility and supports a wide range of data structures. However, no demand profile or generation output was provided, as they have to be supplied by the user. There is also a limited range of libraries, as it only focuses on energy system components, meaning it lacks a comprehensive set of pre-configured modules tailored to industrial process heating technologies. Thus, it would be limited to designing an extensive heat system configuration.

TRNSYS is proficient in dealing with modelling systems that require detailed temporal resolution, as it could generate dynamic load profiles with high granularity over a time span (Narayanan *et al.*, 2020). TRNSYS enables a degree of flexibility when importing external data to the tool. The software includes an extensive library of modular components which can be deployed to the energy system process, in addition to customisable unit operations that include heating technologies (Jani *et al.*, 2020). Therefore, TRNSYS is suited for complex and time-varying industrial energy use patterns.

Table 2.6: Specifications on constructing an energy modelling framework (Lahijani, Protheroe and Gschwendtner, 2025).

Basis	**Description**
Location-specific modelling	The energy modelling tool should be equipped with functionality to capture site-specific parameters such as weather data, capacity factor, distance between infrastructural assets, proximity-based resource and cost concerns, industrial processes and other supply chain considerations. This function is important in the energy model for assessing industrial process heat, which is unique between each site due to different processes and sector types.
Capacity sizing	This feature often deals with optimising the sizing of technological equipment. This feature is particularly important for energy models to evaluate the techno-economic feasibility of a novel alternative technology and to understand the implications of the setup without actually deploying the technology in real life. The sizing allows assessment of the process heat system to be most cost-efficient and reliable.
Technology selection	The technology pathway is an essential criterion for the energy modelling tool to be efficient. The technology selection precedes capacity sizing, as it is used to evaluate fuel-switching tools and emerging operations which are important to support the decarbonisation efforts.

Table 2.6 (continued): Specifications on constructing an energy modelling framework (Lahijani, Protheroe and Gschwendtner, 2025).

Basis	Description
Multiperiod modelling and projections	Dynamic simulation helps to simulate models to understand the time-varying parameters in the supply and demand concept. The demand response, especially heat load, which is affected by the ramp-up of production in the factories, helps to create a more robust prediction of the model, ensuring reliability for energy planning. The need for projection is to create a hypothetical condition on how the system configuration would look like, which supports the claim based on the assumptions made in the model.
Economic Analysis	Economic assessment should be easily employed in the modular component where the cost function could be coupled along with the energy transformation in the component. Therefore, the decision not only could be compared with alternate technologies, using investment costs, but also could be compared along with the cost for the throughput of the energy flow, with a better calculation for cost per unit volume instead of just an arbitrary cost.
Scenario Analysis	The model should be capable of incorporating various changes in parameters, such as fluctuations in energy prices, energy availability, market conditions, and policy frameworks. Sensitivity analysis could also be simulated to show if the trend of technological pathways is consistent or succumbed to slight deviation.

EnergyPLAN is designed to conduct macroscale-level analysis; hence, it would lack details on representing industrial processes. It is widely used due to the high-level strategic assessment while having a fairly good level of capability in handling large datasets (Noorollahi *et al.*, 2021). Although Calliope's input has to be relied on externally by the users, it is flexible in managing a wide range of data structures and formats, including YAML and CSV files (Wang *et al.*, 2023). As a result, Calliope could simulate data down to the hourly or sub-hourly time scale only when the detailed data of the temporal resolution were provided.

In summary, despite the advantages provided by the energy modelling tools reviewed, these tools are not designed to fully achieve the research aims of this thesis. Although they include industrial heating applications, none of the tools support features such as temperature cascading, efficiency, retrofitted boiler technologies, fuel-switching, and integrated modelling of electricity grids and bioenergy networks all at the same time. Moreover, scalability remains a key challenge. Modelling all major industrial sites, electricity generation facilities, and biomass resource nodes not only increases computational runtime but also makes the interpretation of network connections more difficult.

For this reason, a generic tool built around the specific problem is required. Such a tool can remain lean, scale only what is necessary, and provide an appropriate level of interface while maintaining transparency. It must explicitly represent process logic, including energy flows and transformations, and therefore needs to be developed using a generic modelling language.

2.6 Optimisation and Modelling Frameworks

Following the review of industrial process heat and energy system modelling approaches, this section examines the modelling and optimisation frameworks relevant to the development of the thesis model. Since existing energy system tools do not fully support the required level of site-level industrial representation, technology selection, and integrated electricity and bioenergy modelling, a generic modelling language is required. In this context, Pyomo, P-graph, and mathematical programming methods provide the basis for developing a modular and scalable modelling framework.

The literature on Pyomo is important for outlining the progress of the Pyomo algebraic modelling language, which is open source and provides a good starting point for moving away from proprietary licensed energy modelling tools, which are also typically black-boxed. In addition, Pyomo is a mature software package that is compatible with different kinds of solvers, which can be used interchangeably when solving the problem. On the other hand, P-graph Studio and its Python software package are also open source, with an available interface in P-graph Studio that provides a good visual representation of the superstructural model in a Process Network Synthesis manner. Together with Pyomo, mathematical programming is discussed to identify the suitable programming method to be included as solution algorithms in the energy modelling framework.

2.6.1 Generic modelling language

Pyomo Algebraic Modelling

Pyomo is a collection of Python software packages which solves optimisation problems in Python with a wide range of mathematical formulations. Pyomo supports an object-oriented programming method which includes a diverse set of components such as sets, scalars, decision variables, parameters, objectives, constraints, equations and many more. The optimisation algebraic modelling language is open-source and is comparable to GAMS, GurobiPy, (A Mathematical Programming Language) AMPL and many other commercial programming language models.

Many modules and toolkits have been developed and integrated within the Pyomo library designed for their specific purposes. Some notable extensible ecosystems from the Pyomo library include SolverStudio and TEMOA. Other main energy systems models are developed, such as PyPSA, Calliope and ERGET. Pyomo is a stable and extended Python optimisation library which has wide applications from energy modelling, process and industrial systems, research and education, as well as equipment scheduling and planning. For example, Pyomo has been applied in tools developed to study operational demand and sizing optimisation using Mixed-Integer Nonlinear Programming (MINLP) formulations for large irrigation systems (Costa Dilmé *et al.*, 2024). It is also noted that the development tool is derived from a Python-based Pyomo library, which further highlights the importance of open-source tools in dealing with complex modelling environments (Costa Dilmé *et al.*, 2024).

In the energy system domain, several studies have been demonstrated on the usage of Pyomo with MILP for sizing capacities of components and operational optimisation of electricity generation units to satisfy consumer demand (Roth, Franke and Rinderknecht, 2023). Nevertheless, the use of Generalised Disjunctive Programming (GDP) is more applicable for comprehensive energy modelling systems. In the thesis, the model not only intends to perform operational optimisation and capacity sizing but also function as a high-level modelling framework that captures logical conditions and discrete technology selections. In the effort of analysing and optimising multistage mixed refrigerant systems, Roth, Franke and Rinderknecht (2023) concluded that the use of GDP is advantageous to address issues associated with non-linearity and non-convex problems from a computational standpoint. This conclusion is because GDP formulations represent logical conditions in a robust manner and improve convergence behaviour during optimisation.

Pyomo.GDP is a modelling extension which enables users to apply mathematical formulation in the general disjunctive space from the operations research community. This is particularly suited for the instances where there is a disjunction choice between two or more operations (Bernal *et al.*, 2022). In the GDP framework, discrete decisions are adjusted for the selection of one or more operations where inequalities and equations will enforce a constraint so that the specified constraints will not be in conflict with each other, thus avoiding any infeasible behaviour within the solution. In the GDP formulation, discrete decision variables determine which configuration is selected, whereas the associated equations and inequalities enforce relevant constraints based on the configuration made. Moreover, the Integrated Design, Evaluation, and Analysis of Systems (IDEAS) platform has developed a Generation and Transmission Expansion Planning (GTEP) package based on Pyomo (Skolfield *et al.*, 2025). The IDAES framework implements a modular and flexible GDP formulation specifically designed for infrastructure planning problems (Skolfield *et al.*, 2025). The model formulation focuses on decision-making related to infrastructure investment, allowing the system to evaluate failures, disruptions, and grid resilience within long-term expansion planning studies.

2.6.2 Structural Representation of Process Networks (P-graph)

In addition to the algebraic modelling environment, the structural representation of process systems is also important for the development of the thesis framework. P-graph represents a process system as a directed graph, where operating units are depicted as nodes and the connections between them are represented by arcs. From the structural aspect of the P-graph framework, it consists of six primary sets: M, P, R, O, α, β. Set M contains material nodes where the elements are preconditions and consequences that appear in the system structure. In the set representing R, it is comprised of resource nodes, which is the preconditions, whereas set P represented by product nodes and consequences. Set P also consists of intermediate product nodes, which serves as materials that function as consequences for upstream processes and as preconditions for downstream nodes within the network. Operating units represent activities or processes within a system, and both operating units and material nodes are clearly disjoint (Equation (2.7)) (Friedler, Orosz and Pimentel, 2022).

$$M \cap O = \emptyset \quad (2.7)$$

The resource and product nodes are included in set M.

$$R \subseteq M, P \subseteq M \quad (2.8)$$

Since the set of resources represents the precedents and the set of products represents the consequents, an element cannot simultaneously belong to both categories.

$$R \cap P = \emptyset \quad (2.9)$$

The function $\alpha(o_i)$ assigns to each activity the set of materials or conditions that must be available before each process $o_i \in O$ to occur.

$$\alpha: O \to \wp(M) \quad (2.10)$$

In the P-graph framework, operating units $o_i \epsilon O$ have a corresponding set of consequence material nodes assigned to them which is represented by the function β(o_i). The function maps each operating unit to the subset of materials it produces downstream. The function (Equation (2.12)) can be expressed formally as:

$$\beta: O \to \wp(M) \quad (2.11)$$

As shown, $\wp(M)$ denotes the set of M, such that all possible subsets of the material nodes. Consequently, from the expressions in (2.10) and (2.11), each operating unit can be expressed as the mathematical notation:

$$\forall o_i \in O: (\alpha(o_i) \cup \beta(o_i)) \subseteq M \quad (2.12)$$

The concept of the framework consists of three fundamentals: the representation of PNS problems with P-graph; five axioms to be satisfied for in order to solve a process network problem, the foundations of combinatorially feasible process-networks, equipped with three algorithms (Heckl *et al.*, 2015). These algorithms are Maximum-Structure Generation (MSG), Solution Structure Generation (SSG), and the Accelerated Branch and Bound solver (ABB) (Heckl *et al.*, 2015). Figure 2.12 demonstrates the representation of a PNS problem in the P-graph framework which satisfies the axioms.

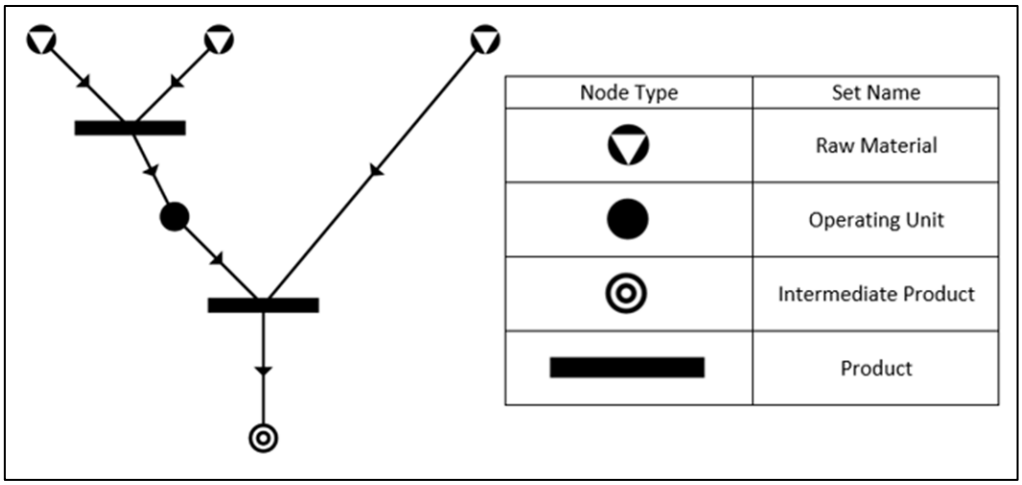


Figure 2.12: P-graph representation of a process network problem.

The axioms are stated as follows (Friedler *et al.*, 1992):

1. All product nodes must be linked with an arc in the structure.
2. A raw material is an element of set M, that is not a consequent of any operating unit node.
3. All operating unit nodes are present in the superstructure which are specified in the problem.
4. Every product node must be connected to an operating unit.
5. All material nodes should not be independent such that they must be linked to at least one operating unit in the structure.

The use of P-graph presents multiple benefits. P-graph is an open-sourced tool that presents itself with a graphical interface, suited for users who prefer an intuitive and unambiguous application framework. This means that the equations are running at the backend, whereas the syntax is not accessible to the users, as the users would only ensure that the P-graph problem structure is correctly represented by adhering to the axioms.

The property positions P-graph favourably within the interpretable optimisation framework recently proposed by Lumbreras and Ciller (2025), who argue that classical optimisation models possess an inherent, largely untapped potential for transparency, in contrast to black-box machine learning approaches. Lumbreras and Ciller (2025) identify simplification as interpretable optimisation, noting that most existing simplification techniques (e.g., presolving, decomposition) are designed to improve computational efficiency rather than human understanding which often obscure the link between the original formulation and the solution in the process. P-graph's structural representation offers a case where this trade-off is largely avoided. This is because the directed-graph depiction of materials and operating nodes are not an abstract approximation of the problem but functions as a model's native representation which means that the decision-relevant structure remains visible to the user throughout, instead of reconverted afterwards through a separate interpretability layer.

P-graph which is initially a PNS concept which deals with chemical processes, has handled problems such as Pinch Analysis and heat exchanger network optimisation, and has been extended into the energy modelling space. As an example, Aviso et al. (2017) have employed P-graph into assessing generation networks which are cost-effective and reliable during temporal changes. Xu et al. (2023) conducted analysis on the integration of electricity, heat and hydrogen in an energy island system. Xu et al. (2023) employed P-graph to identify near-optimal solutions in the superstructure.

Friedler et al. (2019) have also critically reviewed the properties of P-graph that make it useful in the energy system modelling space, surveying its development and applications while identifying key research gaps and promising directions for future work. Their review highlights P-graph's relevance to process intensification where the modification of processes and equipment to enable more compact, capacity-appropriate designs. Process intensification can involve cascading effects, such as heat cascading, which is directly applicable to studying the techno-economic feasibility of heating technologies within the New Zealand industrial context.

Other research papers have outlined the robustness and efficiency of a visual PNS concept which is available in P-graph Studio to model energy systems operations (Éles, Heckl and Cabezas, 2024). Using the P-graph framework, Éles, Heckl and Cabezas (2024) applied P-graph optimisation to assess thermal-powered generation turbines using biomass-based solid fuel for energy production in a rural region. Éles, Heckl and Cabezas (2024) highlighted the potential of using the P-graph framework to evaluate multiple resources and fuel types which operate different technologies. Lee *et al.* (2020) adopted the P-graph framework within the oleochemical space energy model to assess the feasibility of synthesising hydrogen from palm oil mill effluent across the country for the production of sustainable hydrogen fuel.

Kong et al. (2024) has devised a peer-to-peer (P2P) energy trading plan by implementing the P-graph framework so as to optimise energy management during energy planning. Ji *et al.* (2023) used the P-graph framework to assess the techno-economic viability of producing hydrogen from renewable electricity, highlighting its effectiveness for modelling complex energy conversion pathways.

However, the studies mentioned above, although they adopt macro-level energy modelling frameworks, are largely unidimensional. They tend to focus on a single aspect, such as energy supply or the demand for a single commodity. In addition, the models are typically limited to either a single temporal dimension or a single spatial dimension. This highlights a gap for an energy model that incorporates multiple technological components while integrating geographic information system (GIS) data with algebraic mathematical formulation.

2.6.3 Mathematical Programming

Mathematical optimisation or mathematical programming refers to the process of identifying the most favourable solution from a range of feasible alternatives. In a mathematical context, this process involves selecting values for decision variables that yield the optimal value of an objective function, either minimised or maximised, while satisfying a defined set of constraints (Martí, Sevaux and Sörensen, 2025). The resulting outcome is referred to as the optimal solution of the optimisation problem.

The development of mathematical programming has significantly advanced because of improvements in computing, resulting in the many algorithms available today. Mathematical programming is an important milestone due to the capability of solving complex decision-making problems. The first introduction of mathematical programming was started by George Dantzig, who came out with the linear programming method in World War 2 to plan expenditures and returns in order to minimise expenditures to the army and increase the war expenses to the enemy side. Soon after the war, linear programming method or simplex method, was generalised to either achieve maximisation of profits or minimisation of losses and was later adopted by industries for scheduling and planning purposes (Dantzig, 1951).

Linear programming addresses optimisation problems in which a linear objective function is optimised subject to linear equality and inequality constraints, with the simplex algorithm being one of the earliest practical solution methods (Figure 2.13). Extensions such as non-linear

programming (NLP), mixed-integer linear programming (MILP), and mixed-integer non-linear programming (MINLP) are developed to solve more complex problems with a better reflection on the real problem.

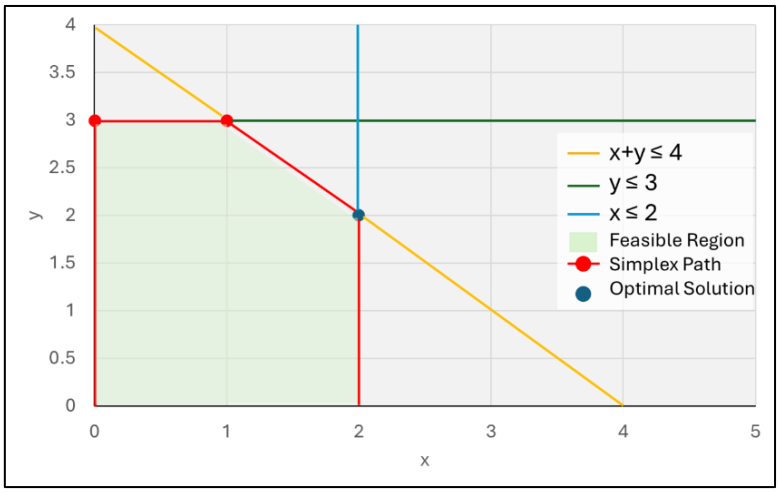


Figure 2.13: Illustration of a simplex method application by solving a linear programming problem with x and y variables.

Optimisation algorithms can be broadly classified into two categories: deterministic and stochastic (Dananjaya *et al.*, 2024). Deterministic algorithms follow predefined rules to identify solutions and do not account for uncertainty within the variable space (Sun, 2013). In contrast, stochastic algorithms rely on probabilistic approaches, using suitable probability distributions to represent and manage uncertainty in the variables (Pedro *et al.*, 2018). A deterministic model will return the same solution all the time when the same input is fed into the model. This makes a deterministic model predictable and consistent. It is widely used in the literature because it provides clear and reproducible outcomes. A deterministic model provides better transparency when conducting scenario modelling, as only specified parameters are altered to observe the change in the solution structure. The transparency also helps modellers to reproduce results and is used for validation purposes.

Non-linear problems

A detailed and realistic representation would typically result in a mixed-integer non-linear programming formulation, as many technical characteristics that constrain operational flexibility or component selection are often simplified or excluded (Kotzur *et al.*, 2021). Mixed-integer non-linear programming (MINLP) problems are often complex, as they merge both the difficulties of combinatorial effects to solve discrete variable sets as well as the challenge of dealing with non-linearity simultaneously. Mixed-integer non-linear programming encompasses both non-linear programming and mixed-integer linear programming as special cases.

Additionally, some problems would introduce non-convexities when undertaking NLP problems or MINLP problems (Tawarmalani and Sahinidis, 2002). Non-linear equations $h(x) = 0$ would often exhibit non-convexities unless they can be rewritten in a convex form. If either the objective function or the constraints are non-convex, the optimisation problem may have multiple local

minima instead of a single best solution. Figure 2.14 visualises the difference between a convex and a non-convex problem when non-linearity is in the problem. General methods to tackle non-convexities are spatial branch and bound techniques, convex relaxation and random initialisation, often generated at random points (Belotti *et al.*, 2013).

However, although these strategies help to solve non-convex problems, it would be more efficient to eliminate non-convexities in constraints by not having them in the first place. This means that the equations introduced are linear. Supported by multiple literature sources, the design of large-scale energy system modelling usually uses MILP, as the key important goal is to search for the most economically feasible pathway, therefore, computational space should be mostly allocated to solve binary decision variables. Also in certain cases, depending on the goal of the research, the LP method could also be employed for simple models with aggregated technologies, where on/off operations do not affect the study. Table 2.7 outlines the guidelines of linearising for non-linear and continuous constraints in the energy modelling systems.

Table 2.7: Strategies of linearisation methods for non-linear and non-continuous constraints which exist in energy system models.

Type	**Example**	**Linearisation Method**	**Resulting Programme**	**Implications on Complexity**
Continuous	Specific investment costs	Binary Steps / Special Order Sets Type 1	MILP	High: Increase the number of binary variables (Milan *et al.*, 2015)
		Piecewise Function / Special Order Sets Type 2	MILP	High: increase the number of binary variables (Belotti, 2013; Milan *et al.*, 2015)
		Intercept slope	MILP	Medium: increase the number of binary variables (Hohmann, Evins and Lygeros, 2017)
		Constant	LP	Low: Non-linearities are omitted without addition of new variables (Al-Wakeel, Wu and Jenkins, 2017)
Non-continuous	Minimum loads	Big-M	MILP	Medium: Increase the number of binary variables (Carrion and Arroyo, 2006)
Nested function	Efficiency depends on the load and unit size	Simplification: One representative variable is used	MILP	Low: Substitute non-linear relationships with piecewise functions and eliminate two variables to prevent a non-linear program (Milan *et al.*, 2015)
Integer-cost component	Total cumulative value of a component	Simplification	MILP	Low: Substitute non-linear relationships with piecewise functions (Heuberger *et al.*, 2017)

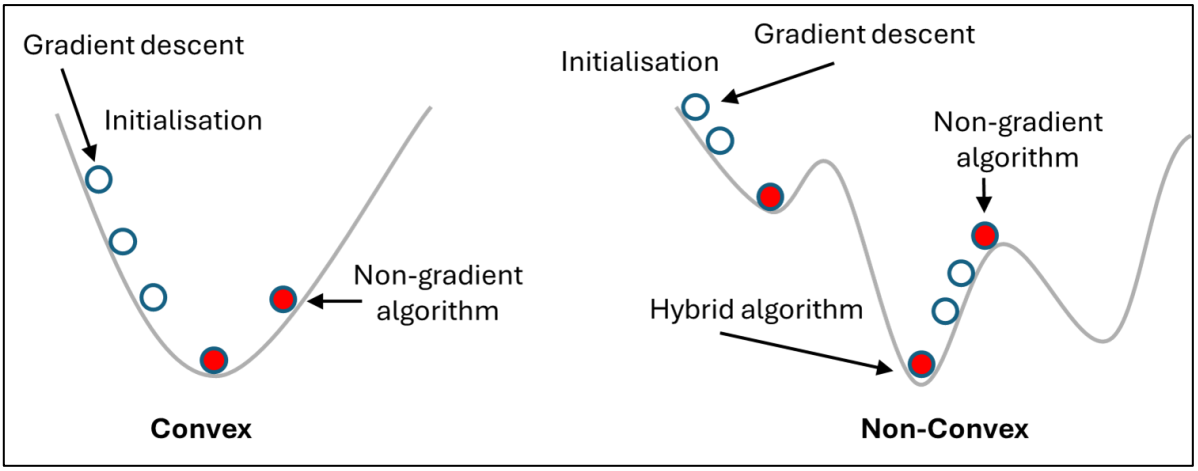


Figure 2.14: The convex and non-convex problems when dealing with a non-linearity problem adapted from Zhang and Xie (2023).

These are some of the techniques which can be applied to simplify energy system models:

1. Constraints that are continuous non-linear: Various linearisation methods can be used for the reduction of non-linear equations. Linear formulation can be applied as an estimation to a non-linear equation such as linearisation by bound reduction method (Belotti, 2013) or piecewise formulation (Milan *et al.*, 2015) which results in a MILP problem. For example, investment cost, which reduces in cost per unit due to economies of scale.
2. Discontinuous constraints: Constraints which include mutual exclusions, such as either–or conditions, if–then–else logic, absolute values or minimum operators, can be represented in linear optimisation models by using the Big-M formulation and convex hull formulation. The Big-M formulation approach introduces binary variables together with a sufficiently large constant to enforce logical relationships, which naturally results in a mixed-integer linear programming formulation. The convex hull approach, on the other hand, reformulates the disjunction into a set of linear constraints which describe the convex hull with a tighter feasibility region.
3. Multivariate relationships: Functions of the form $f(x_1, x_2)$, such as efficiencies that depend on the mix of biomass resource type can be represented in a linear framework by fixing all but one variable at predefined values and precomputing the corresponding outcomes (Milan *et al.*, 2015).
4. Multiplication of a binary variable and continuous variable: Terms involving the multiplication of a continuous variable and an integer variable can be reformulated by introducing an auxiliary continuous variable, with the original variables recovered during post-processing. For example, binary unit-level on and off decisions is replaced with integer variables per technology type and integer scheduling variables with continuous variables in some runs (Heuberger *et al.*, 2017).

Simplification is required to reduce model complexity, allowing the problem to be formulated as a mixed-integer linear program or, in some cases, a linear program. Therefore, MILP is widely applied in planning and scheduling contexts, where it is used to determine optimal allocation of

limited resources. As a summary, the literature of the thesis will only focus on MILP and disjunctive programming as part of the scope of the thesis.

Mixed-Integer Linear Programming (MILP)

MILP is generally used in energy modelling for the selection of technology choices (Kotzur *et al.*, 2021). It is often coupled with other algorithm methods to merge time steps without compromising the accuracy of the time series (Stein *et al.*, 2017). The main improvement from LP is that MILP is equipped with integers which give users the option for binary choices. MILP refers to problems where variables consist of integers while some other variables are non-integers. The programming models can be represented as follows (Grossmann, 2021):

$$\min Z = a^T x + b^T y \quad (2.13)$$

$$s.t. Ax + By \leq d \quad (2.14)$$

$$x \geq 0, y \in \{0,1\}^m, \quad (2.15)$$

where x are the continuous variables, and y are the discrete variables which are assumed to be either 0 or 1. Bounded discrete variables could be almost certain to be reformulated in terms of 0 or 1. Some of the simpler ways to solve an MILP model are by using a "brute force approach", in the case where all possible combinations of binary variables y are set as 0 and 1, thereby eliminating the integer features and solve the MILP in an LP format. However, this proved to be computationally heavy.

An alternative method would be relaxing the binary variable so that it would be continuous instead of being discrete, meaning that the value can hover between 0 and 1. The method being effective to reduce MILP to LP, as non-integer solutions are usually obtained from the approach, the relaxed MILP problem would yield a lower bound for the solution. Although rounding solutions obtained from relaxed MILP problems can lead to infeasible or suboptimal results, this limitation motivates the use of the branch and bound method (Grossmann, 2021). This approach systematically explores the solution space and provides feasible solutions that serve as upper bounds on the optimal objective value. The concept of the relaxed MILP method is presented in Figure 2.15.

On the other hand, there are four major methods to solve a MILP problem, which are Gomory cutting planes, Benders decomposition, branch-and-bound search, and lastly, branch-and-cut methods (Achterberg and Wunderling, 2013). The cutting plane method solves a sequence of linear programming problems by repeatedly adding additional constraints that remove infeasible fractional solutions from the relaxed problem (Grossmann, 2021). The newly added constraints progressively restrict the feasible region until an integer solution is obtained. Benders decomposition separates a mixed-integer linear programming problem into two parts: a master problem containing the integer variables and a subproblem containing the continuous variables (Achterberg and Wunderling, 2013). The subproblem generates additional constraints, known as Benders cuts, which are added to the master problem to improve lower bound estimates. This process is iterated until the variables converge and an optimal solution is found.

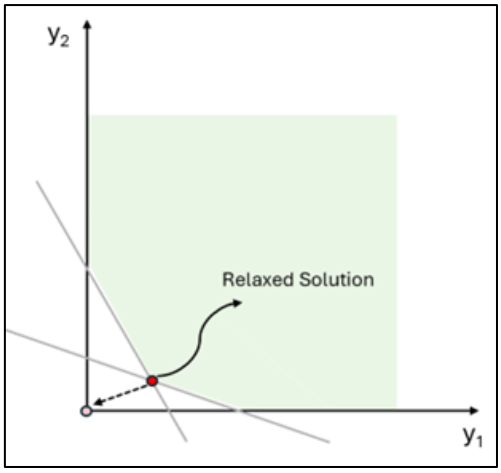


Figure 2.15: Infeasible integer values from relaxed MILP adapted from Grossman (2021).

A branch and bound procedure is applied, where the solution space is explored through a binary tree structure. At each node, linear programming subproblems are solved to generate lower and upper bounds. These bounds are used to discard branches that cannot lead to an optimal solution, reducing the number of combinations that need to be evaluated (Morrison, Sewell and Jacobson, 2014). There are three heuristics in the branch and bound method. In short, the first rule is to order binary variables on the non-integer variables from the relaxed MILP. The algorithm follows this order and looks for the first variable that is not an integer in the solution of the relaxed problem, meaning $y \neq 0, 1$. That variable is then selected for branching. The second rule is to pick a non-integer binary closest to 0.5 to be branched, this means that it would solve integer variables which are "undecided" to speed up the pruning effect (Huang *et al.*, 2021). The last rule is to select the binary variable with the largest degradation. This rule means that the binary variable which is responsible for the largest change in the objective function is addressed so as to progressively identify binary variables which have a greater impact on the objective function (Beale, 1979).

However, despite the ability to parallelise the solving of subproblems or commonly known as threading, there are some limitations to the branch and bound method. Problems could arise in the target to find the solution when the branching strategy is skewed or the number of subproblems is too large (Fukunaga, 2011). This would lead to a long time taken to explore the subproblems, especially when dealing with large-sized problem with intricate constraints and a large search space.

Disjunctive Programming

Many energy system optimisation models formulated as mixed-integer linear programs can be interpreted as linearised representations of disjunctive decision problems. In such models, system behaviour is governed by mutually exclusive operational or investment states, for example, whether a generating unit is committed or not, whether capacity is expanded in a given planning period, or which segment of a piecewise linear cost or efficiency function is active. Disjunctive programming provides a natural conceptual framework for expressing these logical "either–or" relationships, while MILP formulations arise through reformulation techniques such as Big-M or convex-hull representations that enable efficient solutions using commercial optimisation solvers

(Balas, 2018). Binary and integer variables in energy system MILPs therefore encode the underlying disjunctions associated with unit commitment, capacity expansion, technology selection, and technology cost. Figure 2.16 illustrates an example of the application of disjunctive programming in a superstructure model.

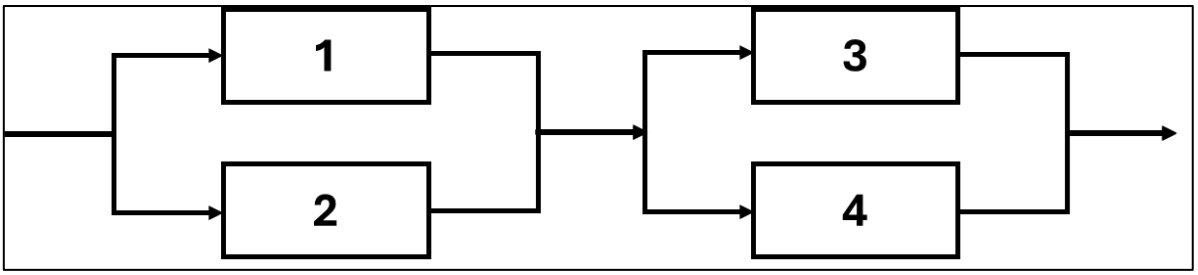


Figure 2.16: Superstructure model with disjunctive operation, adapted from Grossmann (2021).

The main expressions of a general disjunctive programming model which is given by Chen *et al.* (2022), are listed from Equation (2.16) to Equation (2.23):

$$\min obj = f(x,z) \quad (2.16)$$

$$s.t. Ax + Bz \leq d \quad (2.17)$$

$$g(x,z) \leq 0 \quad (2.18)$$

$$\bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ M_{ik}x + N_{ik}z \leq e_{ik} \\ r_{ik}(x,z) \leq 0 \end{bmatrix} \quad \forall k \in K \quad (2.19)$$

$$\Omega(Y) = True \quad (2.20)$$

$$x \in X \subseteq \mathbb{R}^n \quad (2.21)$$

$$Y \in \{True, False\}^p \quad (2.22)$$

$$z \in Z \subseteq \mathbb{Z}^m \quad (2.23)$$

The optimisation problem involves the minimisation of an objective function $f(x,z)$ which is subject to types of constraints. Global linear constraints $Ax + Bz \leq d$ along with non-linear constraints $g(x,z) \leq 0$ apply to the extent of the problem. Conditional linear constraints $M_{ik}x + N_{ik}z \leq e_{ik}$ and non-linear constraints $r_{ik}(x,z) \leq 0$ are categorised into disjuncts and organised into disjunctions indexed by $k \in K$. Every disjunct which is indexed must represent a logical statement where at least one logical statement must be defined as true. The logical proposition, $\Omega(Y) = True$ is an important component in the general disjunctive programming formulations as they represent the logical relationship between the equations and inequalities in the optimisation model. Decision variables can be continuous x, Boolean Y and integer z are also important components in the model.

GDP is proficient in modelling a system where the nature of the decision variables is discrete. The documentation provides its disjunction method in selecting an alternative between processes A

and B in a process design problem. An if-then-else relationship will be expressed as follows in a disjunct behaviour in Equation (2.24). If Boolean Y_1 is True and Y_2 is False, the constraints in the first disjunct will be applied and vice versa when Boolean Y_2 is True and Y_1 is False.

$$\begin{bmatrix} Y_1 \\ constraints \\ for\ then \end{bmatrix} \vee \begin{bmatrix} Y_2 \\ constraints \\ for\ else \end{bmatrix} \quad (2.24)$$

$$Y_1 \vee Y_2 \quad (2.25)$$

In Big-M reformulation, the disjunct is associated with a binary variable y_j which is equal to 1 when the constraints of disjunct j are active and 0 otherwise. The simple expression to describe the disjunct into Big-M reformulation is to relax the constraint when $y_j = 0$ and enforced when $y_j = 1$, as proposed by Raman and Grossmann (1994).

$$A_j x \leq b_j + M_j(1 - y_j) \quad j \in D \quad (2.26)$$

$$\sum_{j \in D} y_j = 1 \quad (2.27)$$

$$y_j = 0, 1, \quad (2.28)$$

where M_j is a large value parameter. The concept of big-M constraints is that the inequality $A_j \leq b_j$ is true when $y_j = 1$. Conversely, the inequality $A_j x \leq b_j + M_j$ will be redundant for M_j if $y_j = 0$. The value for M_j has to be tuned so that it will avoid losing precision and numerical instabilities (Cococcioni and Fiaschi, 2021).

On the other hand, convex-hull reformulation is applied to substitute linear disjunction by convex hull (Raman and Grossmann, 1994). Convex hull of a set of points in a 2D space is the convex polygon that is formed to surround all the points. Convex hull formulations provide a tight representation of a single disjunction, meaning that their continuous relaxation yields the strongest possible lower bound (Vielma, 2015). These formulations are often implemented using extended models, which introduce separate copies of variables for each alternative in the disjunction.

The disjunction with inequalities which are linearised by introducing a variable z_j, to let $z_j = xy_j$, which can be treated as a disaggregated variable x. The overview of the disjunctive formulation is modified into (Chen *et al.*, 2022):

$$x = \sum_{j \in J} z_j \quad (2.29)$$

$$A_j z_j \leq b_j y_j \quad j \in D \quad (2.30)$$

$$\sum_{j \in J} y_i = 1 \quad (2.31)$$

$$0 \leq z_j \leq U_j y_j, \;\; j \in D \quad (2.32)$$

$$y_j = 0, 1 \quad (2.33)$$

It is presented that the convex hull reformulation describes the disjunction of the binary variables by relaxing them to be continuous between 0 and 1. The upper bound on the variables may be unnecessary if the constraints force them to be zero when the corresponding binary variable is inactive. Supported by Figure 2.17, the convex hull approach provides a tight continuous relaxation but uses more variables/constraints, whereas Big-M formulations are simpler but produce weaker relaxations. Convex hull reformulation might pose a potential to slow down solvers (Vielma, 2015).

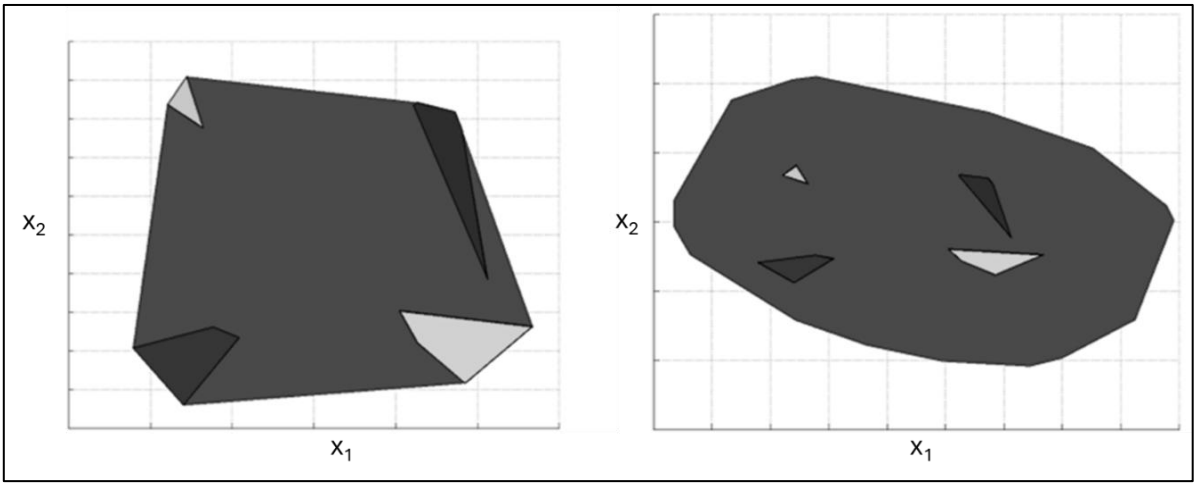


Figure 2.17: Comparison of search space between convex hull and Big-M formulation.

2.6.4 Deterministic Method

Deterministic optimisation methods such as LP, MILP, NLP, and MINLP have been widely applied in energy system modelling for technology selection, capacity expansion, and operational planning. These formulations allow for a detailed representation of system constraints and technology interactions, but as the model size and resolution grow, it becomes harder to compute them. Thus, large-scale deterministic models do often require simplifications to remain solvable, potentially limiting their ability to capture important system characteristics.

To address these challenges, a range of deterministic decomposition and sequential solution methods have been proposed. For example, Benders decomposition, Lagrangean Decomposition and outer approximation separate structural decision variables from more detailed subproblems, allowing the overall problem to be solved iteratively through a sequence of smaller optimisation models. These sequential frameworks improve computational efficiency while preserving model fidelity by progressively refining the cost or feasibility of the solution structures.

In Benders decomposition technique, the large-scale optimisation problem is divided into a master problem and a subproblem that are solved sequentially. The master problem usually contains high-level structural or investment decisions, whereas the subproblem captures better detail on the system behaviour subject to a fixed master decision. The solution from the master problem is inherited, and the subproblem is applied to iteratively refine the master problem until the convergence is achieved (Grossmann, 2021).

There is a broad literature context which uses deterministic sequential optimisation methods to model energy systems. In energy system modelling literature, Benders Decomposition has been applied to separate capacity planning decisions from detailed operational subproblems to solve large-scale expansion problems. García-Muñoz, Dávila and Quezada (2023) introduced Strengthened Benders cuts to reduce the binary variables, which studies the interaction between the peer-to-peer (P2P) energy trading scheme and battery operation. Moreover, studies on capacity planning problems for renewable generation such as hydroelectric, wind and solar are addressed using Bender's decomposition. Hence, research on using Benders' decomposition for discrete variable problems is currently expanding (Zou, Ahmed and Sun, 2018).

Building on this literature, one of the solution algorithms is inspired by Benders Decomposition. The solution algorithm developed not only applies single-stage deterministic methods but also adopts a sequential deterministic optimisation approach in which high-level fuel-switching decisions are optimised using simplified cost representations, followed by detailed evaluation of upgrade costs in a secondary model, as shown in Figure 2.18. The two stages are solved iteratively until cost convergence is achieved, enabling scalability to large-scale energy system models while maintaining computational efficiency.

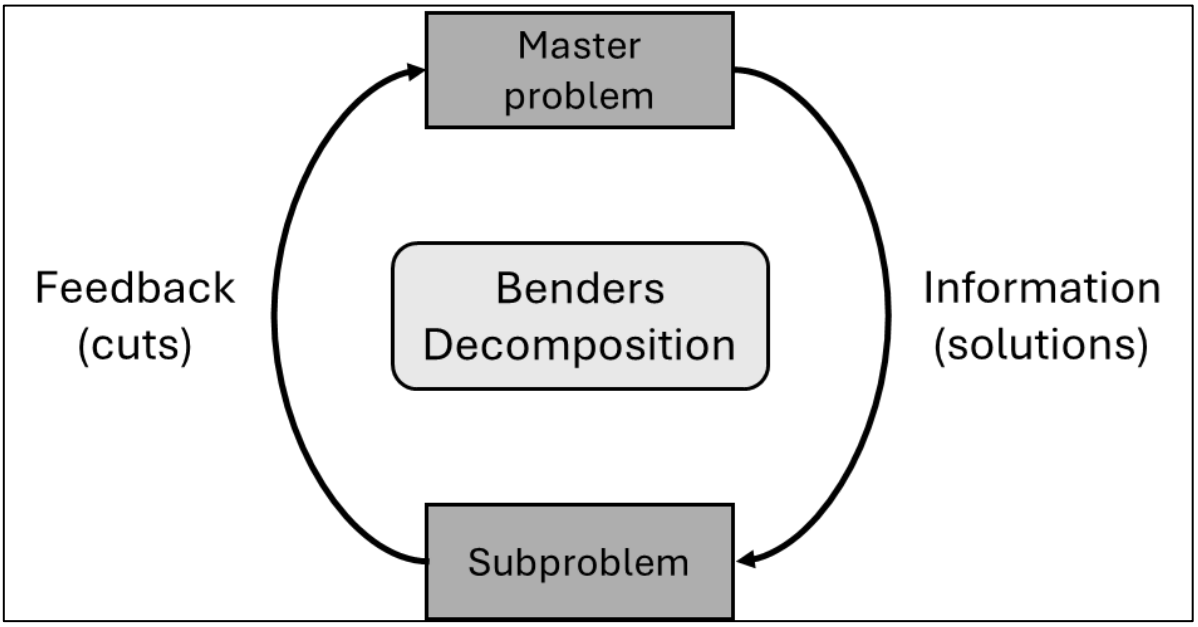


Figure 2.18: Schematic of the Benders decomposition method, adapted from Rahmaniani et al. (2017).

2.6.5 Monte Carlo Uncertainty Analysis

Stochastic programming is a type of solution algorithm which deals with unpredictability in the problem. The algorithm would yield different results for each simulation run despite having an identical initial condition because it is random in nature. Due to the extent of the literature, Monte Carlo uncertainty analysis will be the focus in the chapter. Other stochastic methods are not applicable due to the nature of the energy system models in general. MILP models are linear and solution methods such as branch and bound or cutting planes are far more suitable than genetic algorithm, particle swarm optimisation and simulated annealing methods, which solve problems which are non-linear or non-convex. However, the Monte Carlo method is useful, as it is not a solution algorithm but a stochastic sampling technique (Kissell, 2020). Its main use is to apply repeated random sampling to model the probability of different outcomes in a complex problem.

The probability distribution of uncertain inputs and parameters is recorded and then run through the model to understand the underlying risks and outcomes in the problem. Instead of a single deterministic estimate, the Monte Carlo method will provide a statistical foundation for assessing the range and probability of possible outcomes (Al-Hanoot *et al.*, 2025). In essence, the mean and standard deviation of a parameter or an input variable are required for the model. The expected mean of the variable or parameter is estimated by averaging all the datasets (Villemin *et al.*, 2024).

$$X = \frac{1}{N}\sum_{i=1}^{N} X_i \qquad (2.34)$$

where X is the mean value, N is the total number of sample and X_i is the value of the inputs or parameters at sample i.

The variance and standard deviation are calculated to quantify the difference within the dataset. These statistical measures are important to study the uncertainty of the results. The equations for variance and standard deviation are as follows (Al-Hanoot *et al.*, 2025):

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N} (X_i - X) \qquad (2.35)$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (X_i - X)} \qquad (2.36)$$

where σ^2 represents the variance and σ denotes the standard deviation.

The Monte Carlo method has been applied in many energy system modelling applications. Al-Hanoot *et al.* (2025) employed the Monte Carlo method to solve probabilistic optimal power flow problems, which reduces the cost by up to 9%. It is also applied in microgrids for better energy management for renewable resources (García Vera, Dufo-López and Bernal-Agustín, 2019). As a summary, stochastic methods are important in real-world energy modelling applications,

especially when dealing with uncertainties or to test the robustness of the model structure. Despite randomness in nature, it creates a flexible approach which could be adopted across a wide range of engineering operations. In addition, Shah (2025) applied the Monte Carlo approach to quantify renewable energy planning in the USA. Shah (2025) put a strong emphasis on the Monte Carlo uncertainty analysis helping improve efficiency and prediction accuracy by 15-25%, making the renewable energy strategies more robust.

As a summary for the section, the optimisation environment and formulation choices relevant to building a scalable backend framework which is capable of handling large problems are critically discussed. Mathematical programming methods, along with the strategies to reduce complexity in the mathematical framework have been outlined in the literature not only to build a problem of a magnitude but also to be able to solve the problem without any infeasibility.

2.7 Conclusions

From metallurgical to chemical, manufacturing to dairy processing, the dizzying array of processes that represent the industrial sector makes decarbonising individual sites independently a daunting task (Thiel and Stark, 2021). This chapter introduces industrial process heat systems and reviews heating technologies relevant to fuel-switching. The challenges considered extend beyond economic costs by also including spatial factors related to the location of resource supply and demand centres. The literature also emphasises the importance of categorising process heat according to temperature requirements, which highlights that appropriate classification is essential for achieving different modelling objectives.

While there are many existing energy modelling tools available in the market, none of the tools could integrate site-level details with high-level macroscale energy modelling simultaneously. Scalability also remains a key limitation, as it is difficult to determine which components should be scaled and which should remain fixed, often resulting in inefficient model frameworks. Nevertheless, a broad context of literature has been examined as a guideline to simplify a complex energy modelling network. Generic and open-source modelling tools such as Python and P-graph are explored, where Python provides the mathematical formulation and automation, whereas P-graph provides the conceptual framework on how combinatorial problems work through PNS. P-graph-based software, which is also open access, provides a visual representation of the problem to a certain extent.

In addition, this chapter also provides the fundamentals of mathematical programming, including MILP and NLP, as well as GDP. These mathematical programmes are discussed so as to understand the nature of each program and which program is suitable based on the model scale and complexity. To improve the performance of the model and solution reliability, deterministic solution algorithms with the sequential optimisation concept are introduced. Stochastic methods such as Monte Carlo are adopted in the energy modelling tool to address uncertainties in the energy modelling framework.

2.7.1 Gaps in the Literature and Key Advancements of the Thesis

Gaps in the Literature

Based on the literature reviewed, there are several key research and energy modelling gaps which motivate the development of the energy modelling tool in this thesis. Existing studies on process heat often focus on plant-boundary improvements such as heat recovery, process integration, or individual fuel-switching options, while broader system constraints such as electricity supply, biomass availability, and delivery infrastructure are frequently treated as out of scope. Conversely, energy system models are often useful for identifying high-level transition trends, but they typically rely on aggregated representations that do not adequately capture site-level industrial process heat demand, temperature quality, or structural system configurations under infrastructure constraints.

In addition, there is limited capability in existing modelling tools to represent customisable modular operations while simultaneously integrating multiple interacting layers such as industrial process heat technologies, electricity generation expansion, transmission and distribution constraints, and biomass logistics. Although algebraic modelling environments provide flexibility, they often have a steep learning curve and may unintentionally alter the mathematical structure of the problem. P-graph provides intuitive structural representation, but scalability and solver accessibility remain practical limitations for large-scale applications. As a result, current approaches do not adequately support site-level, spatially explicit, and structurally optimised modelling of industrial process heat systems at a national scale.

This thesis develops a process heat-centric, spatially explicit, and modular energy system modelling framework that integrates temperature-resolved industrial demand with electricity grid and bioenergy network interactions, while supporting technology selection, capacity sizing, and scalable structural optimisation for national-scale analysis. In essence, the detailed research gaps are outlined as the following:

A. Limited integration of spatial modelling with process heat quality constraints

Many energy system models adopt a top-down approach with aggregated representations. These are useful for identifying high-level trends, but they often treat the system as a black box, thereby limiting visibility into underlying technological and operational details. Moreover, structural system configurations and network constraints, which can strongly influence outcomes through cascading effects, are not explicitly represented. As a result, modelling approaches that provide higher spatial resolution often do not capture process heat quality, particularly temperature grades and heat cascading within industrial sites.

B. Limited capability for customisable modular operations in existing tools

Developing technological blocks is often difficult for external modellers. Algebraic modelling environments provide flexibility but have a steep learning curve, and implementation choices can unintentionally change the mathematical structure of the problem, for examples from MILP to MINLP. P-graph provides an intuitive way to define modular operating units, but scaling models

to large systems and reproducing them consistently remains challenging, particularly when many repeated structures are required.

C. Limited structural-level optimisation for large-scale, multi-layer energy systems

Existing tools often prioritise outcome projections over structural optimisation. In particular, there is limited capability to optimise system structure while simultaneously representing multiple interacting layers, such as industrial process heat technologies, electricity generation expansion, transmission and distribution constraints, and biomass logistics. This limits the ability to identify cost-optimal configurations under realistic infrastructure constraints.

D. Limited solver accessibility and flexibility for large-scale P-graph implementations

Although P-graph offers strong structural clarity, solver accessibility is a practical limitation. Large problems may exceed the capability of built-in or default solvers, and solver options are often restricted or difficult to configure. This limits scalability and constrains computational performance. A standalone P-graph is also difficult for accessing simulation modes such as sequential optimisation and stochastic optimisation when dealing with a large problem.

Key Advancements of the Thesis

Several key advancements have been made as outcomes of this thesis. These contributions are summarised as follows:

A. Development of an industrial process heat-focused energy modelling framework

A novel energy modelling framework is developed with a specific focus on industrial process heat decarbonisation. This framework includes capacity sizing with temporal dimension and enables sizing based on integer unit specifications. A contribution of the work includes modularising energy technologies while preserving the specifications of individual components. To the author's knowledge, this is the first framework that is process-heat-focused as well as integrating detailed process heat-focused modelling with grid reinforcement analysis.

B. Development of a highly spatial energy modelling framework.

The modelling framework incorporates high spatial resolution alongside temporal representation. These granularities enable detailed analysis of energy transformation, upgrades, and regional resource distribution. This allows for improved superstructural representation of real-world system behaviour across locations and time.

C. Integration of the PNS concept within the Pyomo optimisation framework.

This work presents the integration of P-graph visual representation with the Pyomo solving framework through a dedicated porting tool. This enables the use of advanced and more powerful solvers beyond the native P-graph environment. The integration overcomes the limited solver capabilities of the P-graph environment, while providing an interface for users that is not available in traditional mathematical formulations.

D. Integration of geospatial analysis with energy system modelling

Geospatial analysis tools, including ArcGIS and Python-based ArcPy, are integrated within the P-graph and Pyomo framework. This enables spatially explicit modelling by linking resources and demands through geographically informed data to be optimised as a problem.

E. Improvement in P-graph usability and result interpretation

The usability of the P-graph framework is enhanced through the development of a structured spreadsheet-based output system. The Excel-based spreadsheet format improves result extraction, analysis, and interpretation by enabling clearer representation of model outputs and solution structures.

F. Expansion of simulation modes within the P-graph framework

The thesis extends the application of P-graph to multiple simulation modes. In particular, a novel sequential (multi-stage) optimisation framework is developed, incorporating numerical root-finding techniques from the SciPy library to resolve convergence between interconnected optimisation stages. While deterministic and Monte Carlo methods have been previously applied, this work represents a significant advancement in applying large-scale sequential optimisation within a P-graph-based energy modelling framework.

G. Extension of a Python-based implementation for renewable generation data acquisition by extending existing R-based tools

A Python-based implementation for accessing the Renewables.ninja API which is initially R-language based, is developed to automate large-scale data extraction and aggregation. This replaces manual workflows and improves efficiency in generating renewable energy profiles for modelling inputs. The following chapter translates the aps identified directly into the framework design choices. This establishes the methodological basis for the contributions outlined in the section.

Chapter 3: Methodological Framework

3.1 Introduction

This chapter establishes the methodological foundation for the thesis by linking the research objectives to the framework architecture and a case study that demonstrates the workflow for energy modellers who may adopt the proposed framework. The general methodology is structured around three key concepts: (1) Process Network Synthesis (PNS) approach, (2) algebraic modelling language, and (3) data preprocessing. Together, these concepts provide the building blocks for developing the energy modelling framework.

The research questions are addressed by developing a generalised, highly spatial energy modelling tool that focuses on industrial process heat and extends to grid and generation modelling. The framework is designed to be transferable and reproducible for other national or regional settings. This chapter describes the integration of P-graph and Pyomo, including the porting tool that ports P-graph structures into an algebraic optimisation framework. The conceptual design of each energy system component is also presented. A central focus is the usability and functionality, enabling users to import modular operating units for iterative model development and to specify temporal resolution, number of time steps, scenario types, solver choices, and simulation modes for the given study.

To demonstrate the usability of the framework, the P-graph-represented superstructure for key modular components is outlined in the energy model. This enables other users to understand the operations represented by each node within the superstructure and provides the possibility of replicating custom modular components if needed.

To demonstrate the functionality, the methodology is applied to a national-level case study of the New Zealand energy system, with a focus on decarbonising industrial process heat. New Zealand provides a useful test case due to its geographic separation into two main islands and its isolation from international energy networks, which increases the importance of energy sufficiency as fossil fuels are phased out.

3.2 Modelling Concept

The modelling concept presented in this thesis provides a structured framework for developing and implementing a macroscale, highly spatial energy model that supports site-level technology selection for process heat decarbonisation. The framework combines PNS using the P-graph method, which represents combinatorial decision problems using a bipartite graph, with an algebraic modelling language. This integration addresses two common limitations in existing approaches. First, algebraic modelling languages are powerful but often lack an intuitive visual representation of network structure and connectivity. Second, P-graph provides strong structural clarity but can face practical limitations when scaling to large problems, including constraints related to model size and computational resources. By coupling P-graph with an algebraic

modelling environment, the framework retains visual interpretability while also enabling scalable optimisation.

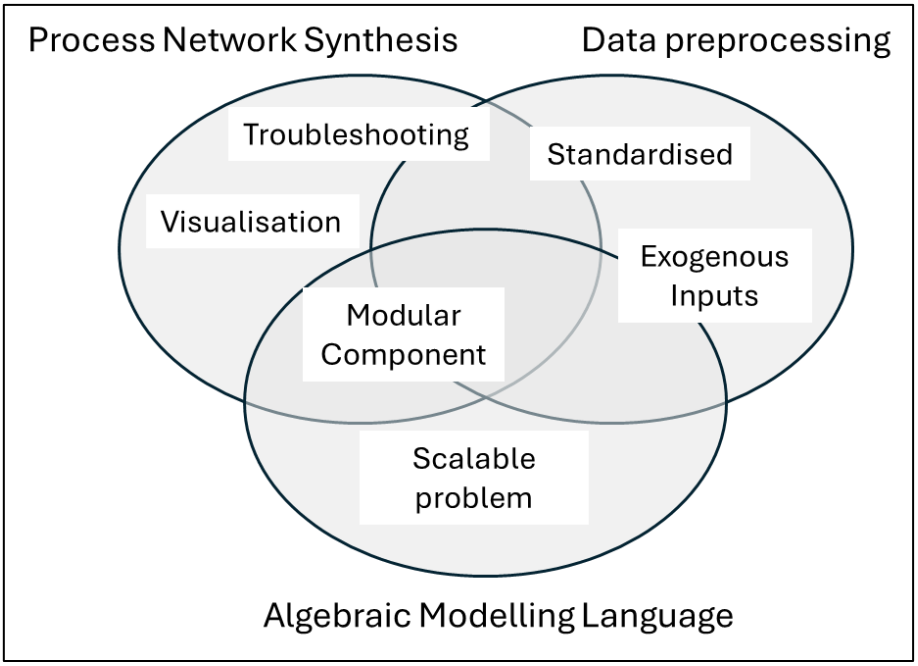


Figure 3.1: Venn diagram illustrating the relationship between the three elements.

The methodology brings together three elements that have traditionally been applied separately in energy modelling.

- Process Network Synthesis - provides a clear structural representation of the energy system as a network of material nodes and operating units. The structure makes precedence and consequence relationships explicit, helps support modular construction of complex systems, and helps identify infeasibilities early by enforcing consistent network logic.
- Algebraic Modelling Language - enables the superstructure to be automatically translated into a mathematical programming formulation that can be solved at larger scale. The framework supports Generalised Disjunctive Programming and can be reformulated into a MILP problem using approaches such as Big M or convex hull reformulation. Constraints and objectives are expressed in a standardised form, improving transparency and reproducibility and allowing access to a wide range of optimisation solvers.
- Data preprocessing - converts datasets into optimisation ready inputs. It includes data compilation, data reconciliation, harmonisation across agencies, and the construction of exogenous parameters needed to instantiate the model. Although it is not part of the optimisation problem itself, it is essential to ensure that model inputs are consistent and suitable for spatial modelling.

The energy modelling framework is organised into six main segments: data preparation, data parameterisation, modular component structures, backend optimisation, model execution modes, and an output layer. Each segment is designed to be generic and flexible so the framework can be instantiated for different national or regional contexts once the required input data are provided.

Model inputs define the datasets and parameters required to construct a valid optimisation problem. Inputs are classified as mandatory when they correspond directly to the components activated in the model structure, and optional when default values or internal libraries can be used. This design reduces configuration errors by enforcing consistency between activated components and the data required to parameterise them. Optional input modules are included within the framework to improve usability, for example by providing standardised technology cost assumptions, default technical constraints, and global model settings that apply unless overridden by case-specific values.

Modular component structures represent the energy system as reusable building blocks. Each module corresponds to a distinct layer within the system and is implemented so that it can be called programmatically within an object-oriented environment. The modules are grouped into two categories. The first category represents energy system utilities, which consume or produce energy, such as industrial process heat technologies and generation assets. The second category represents network infrastructure and assets, which enable the delivery of energy flows through physical or electrical connections. This modular design allows the framework to scale while preserving a consistent representation of technologies and assets across applications.

Backend optimisation enables the model structure to be solved using either a PNS representation or an algebraic optimisation representation. To support large-scale problems and broaden solver compatibility, a dedicated translation mechanism is developed to map the network representation into a Pyomo-based GDP formulation. This module automates the handling of sets, parameters, constraints, and objective definitions in the algebraic environment, reducing the need for user interaction with the underlying mathematical formulation while retaining transparency through the network structure.

Model execution modes provide alternative ways to solve the optimisation problem depending on the analytical objective. The framework supports deterministic optimisation as well as extended execution modes that can be used to manage problem size and represent uncertainty. These modes are implemented as configurable routines that control how the model is solved and how results are recorded across runs.

Output modules support interpretation and post processing of optimisation results. Outputs are organised using unique component identifiers to reduce complexity and enable traceability from results back to model inputs and structural modules. The framework generates structured spreadsheet outputs that capture key model variables and parameters, allowing users to perform customised analysis and visualisation outside the optimisation environment. Export utilities are also included to support optional structural visualisation and basic debugging checks for infeasible configurations.

In summary, the energy modelling framework integrates input management, modular system representation, backend optimisation, execution routines, and structured output generation into a unified workflow. Supporting utilities may be used to assist with data preparation and parameterisation, but these are application dependent and are therefore described in Section B.

The overall framework is structured into two sections which is the core of the energy modelling framework and the implementation of the energy model in a case study. The overall framework can be summarised in Figure 3.2.

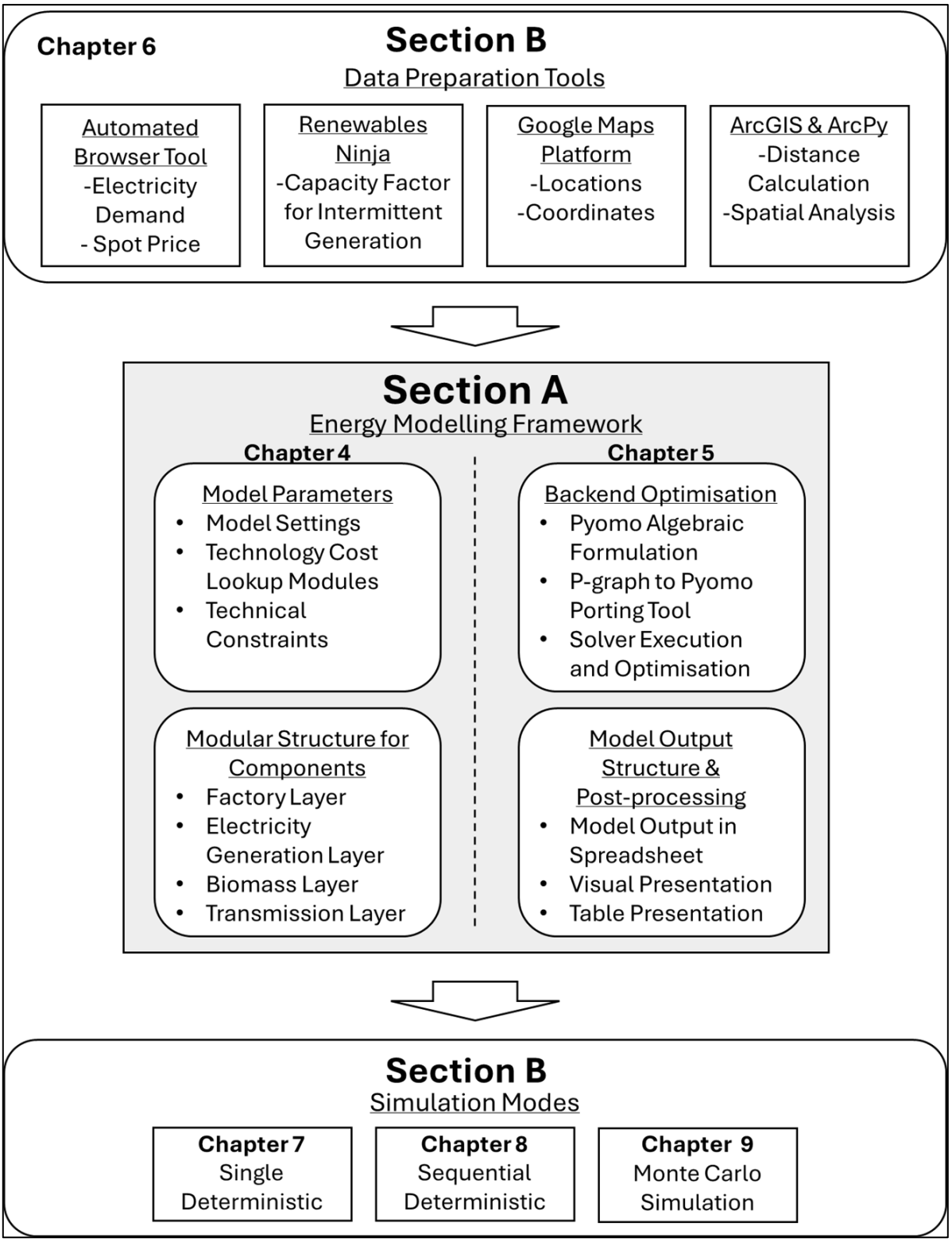


Figure 3.2: Overall methodological framework for the thesis.

Section A: Development of the energy modelling tool

Chapter 4 includes the data pipeline, parameterisation of key technology components that are critical in the process heat energy modelling tool along with their respective technological assumptions. This is followed by the implementation of the technological component inputs discussed earlier as modular operating units within the P-graph based model structure. In Chapter 5 the integration of P-graph with Pyomo and the methodological work on backend optimisation is described.

Section B: Case Study application

This section applies the framework to New Zealand as a national case study. It outlines the practical considerations and specifications required to run a large-scale spatial energy model. The data preparation tools are often conditional, depending on the quality and availability of the data. Although many preprocessing tools are generalisable, they are presented in the context of the New Zealand case study where they are required to generate complete model inputs. The case study is introduced in Chapter 6. Chapters 7 to 9 then describe the simulation modes in applied within the energy modelling framework and subsequently demonstrate the simulation modes supported by the framework.

In summary, these two sections form the backbone of a generalised energy modelling framework that can be reproduced for other national or regional case studies beyond the New Zealand case study context.

3.3 Section A: Framework Development

3.3.1 Motivation

The motivation of Section A is to describe the development of an energy modelling tool that provides fuel-switching pathways for each individual site by minimising total system cost subject to resource and boundary constraints. The model architecture of the energy modelling framework is presented in Figure 3.2. A P-graph visual representation has been used, enabling a gentler learning curve for users which allows flexibility to modify the model or customise unit components without needing to engage directly with the mathematical formulation. This approach forms the backbone methodology of the thesis.

This section focuses on model parameterisation, as it is important for reducing the workload of model users, since most technology blocks are generally consistent in their specifications. Key P-graph modelling techniques are also implemented, including multiperiod time step modelling, technology selection using proportional capacity sizing, and technology selection using integer sizing. Model input parameterisation and the assumptions for technology components, supported by modular operations represented in P-graph, are described in detail to explain both the interactions between nodes within each component and the interactions among components at a higher level.

To solve a comprehensive large-scale energy modelling problem, it involves developing the backend link between P-graph, which provides the visual and structural representation, and the Pyomo algebraic modelling language. More specifically, the P-graph structure is translated and ported into a Pyomo.GDP formulation, which is used to represent discrete decision variables.

3.3.2 Approach

The methodology section in Chapter 4 begins with problem definition and scope. The first section mainly addresses key research questions and literature gaps which were brought up in Chapter 1 and Chapter 2 which would lead the direction of the development of the energy modelling framework. This includes the identification of the objectives of the model, the declaration of variables, boundary constraints and data parameterisation. The problem definition is then followed by a series of broad concepts on how to harmonise both PNS approach and algebraic modelling expressions in the process of developing the energy modelling framework.

The problem definition and scope involve the discovery which ascertain what model input is required for the optimisation model to solve the problem. At the same time, attention on the building blocks of components is outlined which could represent the model architecture, as well as how the components from each layer relate to one another in the energy value chain. Within each modular component, operating units and material nodes are designed to capture key specifications or assigning parameters which are important for results and data analysis.

Unlike mathematical programming where dimensions for layers are defined in sets, the P-graph approach is one dimensional, hence the dimensions are not an important feature due to the initial usability of the approach. Hence, dimensions which segregate components, time space, arcs between layers and nodes within the components are addressed by using the unique identifier for each node and operations.

To improve the functionality and flexibility of the modelling framework, the model can be extended to not only analyse systems model of the factory's perspective but also conduct electricity system analysis. Moreover, uncertainty analysis can also be conducted

The final step of Section A involves an illustrative example of a simple model which comprised of key unit operations developed in the framework.

3.3.3 Summary

Section A outlines the backend development required for the energy modelling framework. However, the application of an actual case study is needed to identify missing preprocessing steps and as a sanity check.

In summary, Section A is the workflow described for the construction of the energy modelling framework whereas Section B is the workflow on how the model is built with the given framework in Section A.

3.4 Section B: Case Study Workflow and Comparative Analysis of Simulation Modes

3.4.1 Motivation

The content of Section B refers to the application of a case study and the development of a generalised preprocessing tool for generation for exogenous model inputs. The main purpose of this section is to deploy the modelling framework to an actual energy system so as to assess the energy background for sensibility checks and verify the heat technology configurations, electricity grid and biomass connection networks. The purpose of this section is also to demonstrate how nodes and arcs in the P-graph are assigned to the mathematical expressions presented in Section A. This methodology shows how the superstructure can be translated into a large-scale model that can be solved using advanced mathematical solvers. These solvers go beyond the built-in branch-and-bound approach by incorporating cutting planes and other linearisation methods available in the Pyomo optimisation suite.

In order to solve large sized problems, the pre-processing and data preparation are as equally as important as the development of the energy modelling framework due to large amount of datasets representing each component required to be used for the model generation. Figure 3.2 presented in this section would outline the necessary workflow required to employ the energy modelling framework, showing a wide range of settings to be defined and most importantly the application of simulation modes to solve the problem. To improve the functionality of the framework, simulation modes are demonstrated after the case study has been introduced.

3.4.2 Modelling Approach

The application of the methodology proceeds with deploying the energy model as a national case study using a bottom-up approach. A large-scale model with high spatial fidelity is implemented to demonstrate the usefulness of integrating P-graph with Pyomo. The case study includes 298 major industrial sites, 291 generation locations, and 214 biomass locations connected through 217 point of connections (POCs), reflecting a national scale system with site-level detail. Implementing a model of this size requires both optimisation expertise and a strong understanding of the New Zealand energy system, particularly in relation to fuel-switching decisions.

It is important to derive information and data to represent each component embedded in the system. However, most data were originally designed for a different context, some are also missing, therefore generalised preprocessing tools are developed to add attributes or generate data to be assigned to the model. This includes aggregated energy generation for intermittent renewables, automation for data extraction as well as spatial tools which identify distance calculation or geodetic coordinates. Verification for the preprocessing step is required to show that the tool is reliable and can be trusted in the operation.

After that, once the model and problem table have been defined, the simulation modes will be demonstrated, which are deterministic and stochastic optimisation. This is important because

modellers would encounter several instances where a problem could be solved in a deterministic manner, when the parameters are clearly defined. However, it is always not the case, as there would be uncertainties, especially modelling a real-life scenario, there would be several unpredictable aspects that could not be represented with an arbitrary value.

Moreover, simulation mode in particular, sequential optimisation, inspired by Benders Decomposition method, is applied in Chapter 8 and Chapter 9. This simulation model shows that integrating P-graph with Pyomo can extend the capabilities of mathematical formulation beyond what P-graph alone can achieve. For the case study, a two-stage optimisation iterative approach is applied in the same New Zealand case study setting. The master problem would define the transition pathway for fuel-switching as well as the feasibility of green fuel production (if it is economically viable) with an assumed delivered electricity price whereas the subproblem will calculate the cost and pathway of grid expansion upgrades and electricity generation upgrades required, where finally the cost from the second stage to update the first optimisation stage until the cost of assumed electricity price matches with the upgraded cost.

A quote often attributed to Stefan Pfenninger (2024) states: “Energy system models do not represent natural processes but are assumption laden representations of complex engineered systems, making validation practically impossible.” In this thesis, simulation results are discussed alongside sensibility checks for selected key sites to assess whether the modelled fuel-switching pathways align with industry plans. The results show not only which heat technologies is selected, but also the required grid infrastructure upgrades and the biomass flows, including biomass type, across individual connections.

3.4.3 Summary

Section B describes the application of the energy modelling framework using a real-world energy system that can be replicated for other studies. The case study workflow must be followed carefully to ensure that the energy system components are represented consistently in the optimisation problem. From an industrial process heat perspective, multiple heating technologies, including retrofitting existing boilers, are evaluated for each site while maintaining temperature grade constraints, which lead to different coefficients of performance and efficiency. On the resource side, explicit generation sites are represented, allowing the model to identify the lowest cost designs and locations. The framework also represents intermediate network connections explicitly between sites. The simulation modes further demonstrate the flexibility of the framework, which is a fundamental attribute of a useful energy modelling tool because it provides users with greater freedom in how the model is configured and applied.

3.5 Chapter Summary

The chapter partitioned the methodology section for the thesis into two parts, with the objective of distinguishing the energy modelling framework and case study data processing with comparison of simulation modes. Section A details the data pipeline and the intricate design of the energy modelling framework. Section A also includes P-graph modelling techniques replicate the functionality of an equation-based model without deriving a single mathematical expression

in the work. Together, these components enable the translation of a conceptual energy system into an optimisable mathematical model.

Section A will extend from Chapter 4 to Chapter 5, encompassing two chapters. This chapter presents the modelling framework and defines the scope of the energy system problem it is designed to address. Although the thesis addresses multiple aspects of the energy system, the framework is developed as a unified modelling approach. Section B builds upon the development from Section A to complete the full process required for the fuel-switching and industrial process heat energy transition narrative.

Section B evaluates the framework's ability to represent process flexibility, enabling the quantification of individual energy flow within the system which are difficult to be done in other energy modelling tools. A key challenge lies in accurately quantifying distances between spatial entities across different layers using generalised spatial tools.

Section A: Energy Modelling Framework Creation

In the first section of the thesis, the focus shifts from the methodological framework to the core development of the energy modelling framework. Specifically, Section A describes how a large-scale energy systems model can be constructed and solved efficiently and consistently. This strategy is important for integrating the P-graph modelling concept with an algebraic modelling language. The P-graph modelling concept is supported by an existing software called P-graph Studio, which is equipped with a flowsheet interface that allows users to drag and drop functions to assign and connect material unit nodes and operating unit nodes. This makes it highly intuitive for conceptualising the superstructure and facilitates troubleshooting. On the other hand, the algebraic modelling language is proficient in describing and solving highly complex, large-scale optimisation problems. By leveraging the advantages of both approaches, a high-fidelity energy systems model can be constructed in P-graph and then solved efficiently using algebraic formulation.

Section A begins by presenting the scope and requirements for the development of the energy modelling framework. The purpose is to define the extent of the energy modelling framework developed to address the research aim of the thesis. Layers and components are discussed to understand their purpose and function within the energy modelling framework. After demonstrating several P-graph modelling techniques to define the temporal representation and capacity sizing of technological blocks, key energy components are modularised. These P-graph conceptual methodologies are explained in Chapter 4.

Following the development of modular structures, Chapter 5 explores the backend of the energy modelling framework. The main model structure is described, along with how the energy modelling framework is designed to import data across workflows. First, the conventional P-graph workflow as well as the novel approach are discussed. The P-graph-to-Pyomo Porting tool is designed to convert both P-graph Studio and Python-based P-graph superstructure inputs, although in different formats, streamlining them and parsing them into the Pyomo solving environment. Generic formulations are created to produce a generalised disjunctive programming representation of the P-graph. The workflow for parsing and processing output results is also outlined in Chapter 5, allowing users to have a better experience by reducing the effort required to extract data from the results. An illustrative example is presented to demonstrate the energy model concept before the case study is implemented in Section B.

Chapter 4: Modular Architecture of Components in the Energy Modelling Framework

4.1 Introduction

The assessment of fuel-switching technologies is inherently a multi-criteria decision problem. Multiple pathways can meet future energy demand, and each pathway introduces trade-offs across cost, feasibility, resource availability, infrastructure requirements, and operational constraints. The existence of many candidate pathways increases modelling complexity, but it also enables the model to evaluate and select among a wider set of configurations that can satisfy system requirements at a minimum cost. This chapter defines the modular architecture of the framework; detailed mathematical formulation is deferred to Chapter 5, while case-specific New Zealand inputs are presented in Chapter 6.

The modelling framework incorporates several key components, including electricity generation technologies and installed capacities, the availability and characteristics of biomass resources, emerging green industries, transmission system upgrades, electricity demand from non-process heat users, and process heating technologies and their capacities. Conversion efficiencies and energy losses are characterised explicitly so that system behaviour can be represented consistently across scenarios.

The model adopts a site-specific approach which incorporates individual components for discrete decision-making purposes. As a generic energy model allows users to input values of specifications, it is often unclear how the calculations and equations are used for different technologies. These considerations mark the importance of applying a bipartite graphical network to effectively assess the discrete decision which are designed to be simple and computationally efficient, but also for visual presentation and troubleshooting infeasible problems. The bipartite graphical network is particularly important to present the unit operation within the object-oriented framework, as the approach is suitable for process modelling of complex unit operations to be coherent and be reused.

A central contribution of this work is to modularise energy technologies while maintaining the temporal and spatial features of the individual components. This is important in the later stages as the deployment of these components can be automated in Python, where they can be called and inserted into the flowsheet. Moreover, data parameterisation is important at this stage so as to allow standardisation of components.

The chapter also describes the techniques within the P-graph framework that reduce the size of subproblems and provide systematic means of representing storage, multiperiod optimisation and sizing for capacities of technologies. These features can be achieved with only two primary components, namely Operating Unit nodes and Material nodes. The first part of the chapter will discuss the main components existing in the first stage model, whereas the second half will focus

on the grid asset upgrades. The end of the chapter includes a simple example on how the structure is presented in the P-graph framework.

4.2 Methodological Workflow

A novel method for constructing a highly spatial energy model is presented in this chapter. The main steps required for model construction are outlined in a way that allows users to build, modify, and scale the model with flexibility. In addition to the energy model itself, the chapter also describes the supporting components needed to ensure that the model operates correctly, including input parameterisation, solver configuration, and procedures for presenting results. The complete methodological workflow for the energy modelling framework, Section A is shown in Figure 4.1.

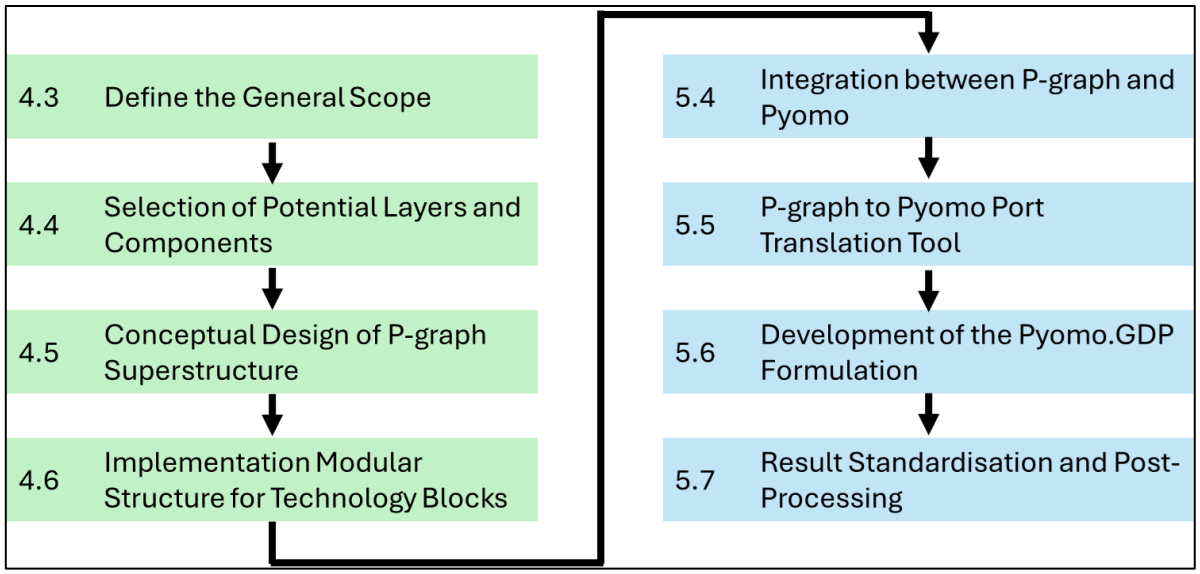


Figure 4.1: Overview of the methodological workflow for Section A.

The workflow begins with system definition, where research questions and system boundaries are specified. These decisions guide the selection of model components and the scope of the case study. Next, technology screening is conducted to identify fuel-switching alternatives that are technically and economically plausible under future scenarios. The screening is informed by the literature and helps to eliminate unrealistic pathways while ensuring that the model boundary remains consistent. As the highly spatial model is built from a bottom-up perspective, site-level information is required for each component.

The modelling technique using the P-graph method is presented after model parameterisation has been discussed. This section demonstrates how the P-graph method with only material nodes and operating unit nodes, can replicate the functionality of mathematical expressions for representing model dimensions such as time, capacity build, integer costing, and proportional costing. This chapter provides the basis for the modular structure section, as these features are implemented in each modular component which represents key elements of the energy model.

This step is explained in detail for selected key components, as it is an important feature for effectively translating the P-graph visual representation framework into a Pyomo-based modelling tool. Modular structures are defined for key components, including heating technologies, generation assets, POCs, and model settings. These modules are stored as reusable structures so they can be called automatically using Python when constructing the model at a chosen temporal resolution. An object-oriented programming approach supports this design by enabling the management of data-intensive structures and automated generation of model configurations.

The optimisation formulation is implemented using Pyomo.GDP to represent discrete technology and network decisions. Technologies are typically modelled as present or absent, and when present, their associated constraints apply. The Pyomo representation generalises sets, parameters, variables, and constraints in a way that reproduces the bipartite structure defined in the P-graph framework. A translation mechanism then maps P-graph structures into the algebraic environment so that the model can be solved using either the P-graph approach or the Pyomo.GDP formulation.

The integration of P-graph and Pyomo is achieved through the construction of a model translation mechanism that maps the P-graph representation into the algebraic modelling environment. This coupling provides the ability to view and analyse complex combinatorial structures through the visual capabilities of P-graph while taking advantage of the efficiency and flexibility of algebraic optimisation within Pyomo. The combined approach overcomes specification limitations that exist in the P-graph framework alone.

In the results and analysis step, both graphical and tabular outputs can be filtered, combined, and organised according to the needs of the user. Results may be grouped or examined based on specific features such as geographical boundaries, time periods, or technology types. These presentations can then be exported for further interpretation, enabling users to draw meaningful conclusions from the case study results.

4.3 General Scope Definition

This section outlines the general scope for Section A's energy modelling framework. The scope in this section elucidates boundaries of the modelling framework and the types of tools designed to represent any application.

The objective of the framework is to model a spatially explicit energy system consisting of industrial demand sites, electricity generation assets, biomass resources, green industries and network connection points. It is designed to capture the interactions between the components using explicit links and flow connections, helpful for representing the energy system as an integrated network rather than as low-resolution aggregated sectors.

The framework supports optimisation decisions such as technology selection, capacity sizing and resource allocation across the modelled system. The framework includes the selection options between process heat technologies, new and existing assets within the system boundaries, and

the fuel selection choice pathway between electricity and biomass delivery, subject to system cost and feasibility constraints. Discrete decisions, in the form of disjuncts, would handle whether the technology or pathway exists, whereas continuous ones represent the capacity sizing and flow volume.

In the energy modelling framework, key inputs are treated as exogenous. This approach reduces the computational burden and allows the optimisation to focus on decisions that are directly relevant to the objective. These inputs include assumptions on coefficients of performance and efficiencies for heating technologies, distances between individual nodes, capacity factors and utilisation factors for electricity generation sites, conversion of biomass fuels from a mass basis to a lower heating value (LHV) basis, and feedstock conversion assumptions for green methanol production in the power-to-X (PtX) pathway. Finally, the framework scope is constrained by design choices that ensure consistent modelling assumptions. As the heating technologies focus mainly on boiler and heat pump systems, the model is intended for indirect process heat applications up till 500˚C where heat is delivered through an intermediate medium such as steam or hot water. Highly process-specific direct heating technologies and very high-temperature applications that require fundamental changes in reaction pathways are outside the structural scope, so they are treated as excluded categories unless explicitly reformulated, customised and added as new modular components as future works.

4.4 Key Layers and Components

The superstructure for the energy modelling framework is presented in Figure 4.2. To answer the research questions on constructing a macroscale energy system model, which is focused on decarbonising industrial process heat, the scope of the study must first be outlined before proceeding to subsequent steps.

At the industrial process heat level, different temperature bands are identified at the factory level. Key heating technologies associated with various electricity and biomass fuel types are included. Biomass resources are allocated across multiple pathways, including thermal electricity generation, gasification for green methanol production, and direct use in industrial process heat.

Electricity generation from different technologies is also considered, with specific attention given to hydroelectric generation with reservoir storage. Electricity can be supplied to POCs for non-process heat electricity demand or allocated to green methanol production via electrolysis of hydrogen. In addition, direct electrification of industrial process heat is incorporated within the model.

Once the general scope is defined, the key layers and components of the energy modelling framework are introduced. This section clarifies how each layer is represented in the model and identifies the core data required for parameterisation. Each component is defined by its role within the energy system, along with its associated attributes, and its interactions with other layers through spatial connections and energy flows.

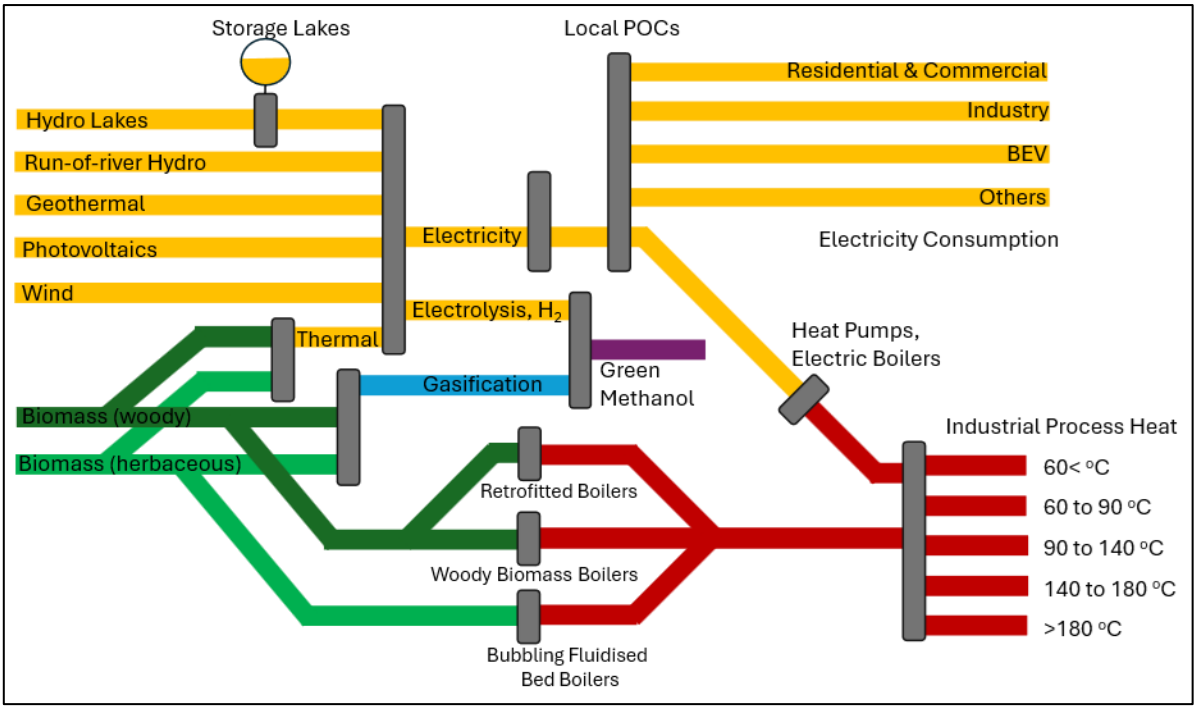


Figure 4.2: Superstructure of the proposed energy modelling framework.

4.4.1 Process Heat

Industrial Heating Technologies

For the thesis, the definition is based on the operating ranges of heating technologies. The primary motivation of this study is to assess the capacity required at the site level to decarbonise industrial process heat, as well as the associated installations needed at the generation and grid levels. Accordingly, process heat demand is characterised not only by temperature but also by defined temperature categories of <60°C, 60 to 90°C, 90 to 140°C, 140 to 180°C, and >180°C. These categories align with the applicability of low temperature, medium-temperature, and high-temperature heat pumps, while electrode boilers and biomass-based boilers are considered suitable for meeting the highest temperature requirements.

The scope also includes replacing fossil fuel-based boiler systems, but industries that require heat loads exceeding 800°C, such as those using furnaces, kilns, or reformers, are not included in this study because suitable low-carbon substitutes for these high-temperature applications are limited. Since heat pumps have different COP values depending on the temperature lift, the efficiencies of the boilers and heat pumps are presented in Table 4.1. The main assumption is that the heat source temperature is the ambient air temperature. This assumption may underestimate the efficiency of the heat pumps, but it is adopted as a conservative approach.

When users do not have site-specific data on process heat demand by temperature range, a default temperature breakdown is applied using temperature shares. These shares are compiled from multiple studies and sources and harmonised into the temperature bands used in this thesis.

Because the industrial sector is highly diverse, the assigned breakdown is selected based on the process category that is most representative of the site shown in Table 4.2.

Table 4.1: Heating technologies introduced within the framework for indirect heating.

Heating Technology	Temperature Range	Fuel	COP/Efficiency	Investment Cost (NZD kW^{-1})
LTHP	<60°C	Electricity	457 %	2000
MTHP	60 to 90°C	Electricity	269 %	2000
HTHP	90 to 140°C	Electricity	173 %	2000
Electrode Boiler	>140°C	Electricity	99 %	1200
Biomass Boiler	All	Woody biomass	80 %	1100
Bubbling fluidised bed	All	Non-woody biomass	75 %	2500
Biomass retrofitted coal boiler	All	Woody biomass	75 %	300

Table 4.2: Default process heat temperature shares by process category, based on the European Commission (2008) and additional referenced and partly confidential industrial data.

Processes	<60°C	60°C-90°C	90°C-140°C	140°C-180°C	>180°C
Asphalt	0	0.05	0.15	0.3	0.5
Bakery	0.05	0.15	0.1	0.3	0.4
Brewery	0.2	0.4	0.4	0	0
Cheese	0.2	0.4	0.3	0.1	0
Chemical	0	0	0	0	1
Dairy Normal	0.29	0.38	0	0.14	0.19
Food Normal	0	0.1	0.1	0.3	0.5
Frozen Food	0.1	0.3	0.4	0.2	0
Meat	0.195	0.305	0.5	0	0
Meat Rendering	0.1	0.1	0.5	0.3	0
Metal	0	0	0	0	1
Milk Powder	0.05	0.1	0.25	0.25	0.35
Milk Powder and Cheese	0.1	0.2	0.2	0.2	0.3
Non-Metal	0	0	0	0	1
Prawn Farm	0.6	0.2	0.2	0	0
Pulp	0.05	0.15	0	0.5	0.3
Salt	0	0	0	0.5	0.5
Sugar	0	0.1	0	0.9	0
Textiles	0.1	0.8	0.1	0	0
Wood Processing	0	0.05	0.95	0	0
Other (high)	0	0	0.1	0.35	0.55
Other (low)	0.2	0.3	0.4	0.1	0

Retrofitting of Coal Boilers

Process heat demand is often attributed to carbon emissions from combusting natural gas, coal, LPG, diesel for heat generation, it provides several challenges to phasing out these boilers. However, coal boilers have proved to be flexible for fuel-switching to biomass resources, as coal briquettes have similar size and density to solid wood biomass derived from logs, wood chips and pellets, so coal boilers can be retrofitted and upgraded into biomass boilers.

The existing primary and secondary fuels in a site have to be outlined, as certain coal and lignite fuel-based boilers can be modified into biomass boilers at 50% of the cost of a new biomass boiler. Therefore, the existing heating infrastructure down to the boiler level for each site has to be addressed when considering the total minimum system cost configuration. The model will provide a more comprehensive yet realistic decarbonisation approach for factories within the region.

Decisions on retrofitting older boilers are often arbitrary and rely heavily on case-by-case assessment. To provide a consistent approach, the default rule in the energy tool was developed based on extensive boiler-level analysis for each site. Coal and lignite boilers with more than 5 MW of installed capacity that were commissioned after the 1990s are assumed to be suitable candidates for retrofitting. This assumption serves as a proxy for scenarios where the high cost of large boilers makes retrofitting more feasible than replacement. In addition, although coal boilers typically have an operational lifespan of 25 to 30 years, their lifespan is often extended through major refurbishment undertaken near the end of their initial operating life.

4.4.2 Dealing with Temporal Variation

In this section, the temporal changes from a reference year will first be assessed and validated before making the projections of electricity demand. Unlike an annual energy model, with a single time step, a multiperiod optimisation energy model has the potential to not only capture the optimised structural configurations but also effectively address the operational problems due to fluctuations in generation and demand. This is particularly important to deal with winter peaks or low renewable energy situations. Therefore, a fuel-switching suggestion on an annual model could be risky when estimating capacity due to oversight of fluctuations, as after accounting for reliability and temporal changes, the technology switch can become costly or infeasible.

Time variation enables load shifting or storage to manage supply and demand imbalance. Multiperiod blocks allow the modeller to assess whether the electrification pushes the system into the capacity limitation rather than assuming electricity is always available. This approach helps the model to improve realism of dispatch and capacity adequacy. Even with a simplified network representation, the multiperiod structure forces generation and demand to balance within each time step, so that high demand periods must be met by available dispatch, storage, or pre-charged energy, and excess generation in lower demand periods can be stored or curtailed. Moreover, the intention of a multiperiod model would help select the amalgamation of electricity generation types, not only covering a single generation type with the lowest cost but also satisfying the demand at each time block.

The same concept applies to biomass use for process heat. This is because the biomass has recoverability, meaning that it takes time to grow before it can be harvested as a bioenergy resource, which would not be accounted for in a single time step model. Therefore, a limit is set on how much biomass resource is available for a time horizon for each time step made. In this chapter, the model is evaluated using a monthly temporal resolution.

The monthly changes are based on the industrial curve, which is the coefficient of utilisation of factories based on their maximum capacity. The assumption is obtained by different sources and considerations. For dairy demand, it is highly dependent on the milk curve, as dairy processing is strongly influenced by the volume of milk that needs to be processed. This sector includes industrial sites which undergo processes such as manufacturing milk powder, cheese, whey, infant formula and ultra-high-temperature milk. Milk production decreases in volume during the winter period. Factors like calving season, availability of pasture and the cattle's' diet affect milk production, leading to dairy industries operating way below their load limit.

Another industry which is widely impacted by seasonal changes is the meat industry. The meat industry includes meat processing, rendering, animal by-products, and a certain level of leather or tannery making. The monthly assumed demand is estimated as a proxy for the weekly slaughter of domesticated livestock in New Zealand (New Zealand Meat Board, 2026). The remaining sectors are less influenced by seasonal changes. Sectors such as metal and food could be analysed by referencing POCs where the demand dominates most of the electricity share, available at every trading period. The metal industry could be observed from Tiwai POC whereas the utilisation factor of the food industry could be represented at a POC in Hastings.

On the other hand, the utilisation factor for wood processing, pulp and non-metal industries is obtained from historical trend plant data. Figure 4.3 illustrates the industrial curve for each sector in New Zealand. The New Zealand industrial load curve is used because the country has substantial economic activity in agriculture, dairy, and meat processing, making it a more representative reference with a larger sample base.

The industry curve will serve as a basis for optimising a multiperiod process heat demand model in Chapter 7, Chapter 8 and Chapter 9. As the process heat and the operability of electricity consumption increase hand-in-hand with one another, the monthly electricity demand for each industrial site can be calculated and aggregated to each POC.

On the other hand, rather than with hourly data, obtaining such data is more challenging because it is difficult to capture high-granularity operational demand data for a large number of factories. This is because operational demand is affected by seasonal variation, day-of-the-week effects, and unscheduled downtime, which can introduce inaccuracies into the model results. Nevertheless, generic load profiles for industrial sectors are obtained from Forschungsgesellschaft für Energiewirtschaft (2021). The load profiles are shown in Figure 4.4. This operational curve will be applied for the four-hourly interval energy model example shown in Chapter 9.

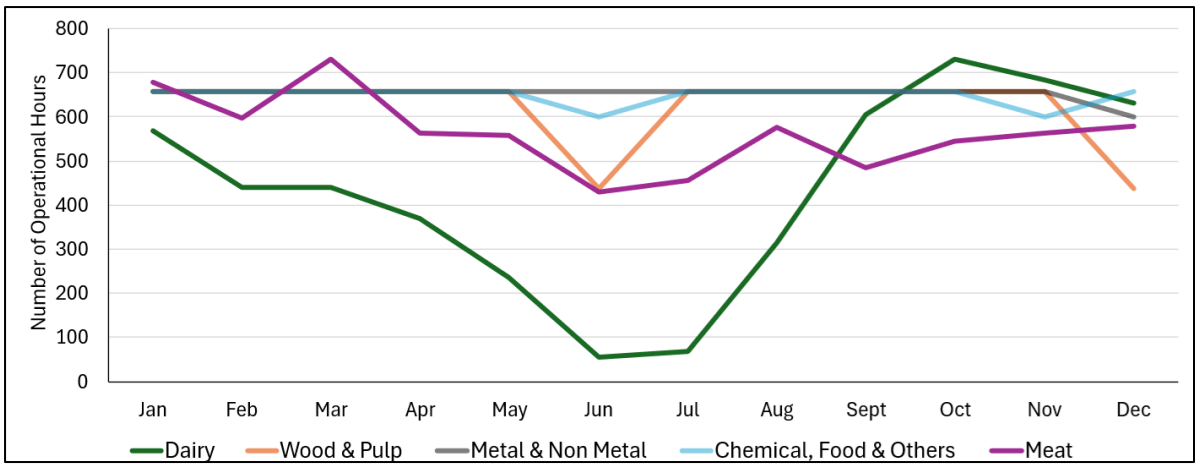


Figure 4.3: Monthly operational hours across different industrial sectors in New Zealand.

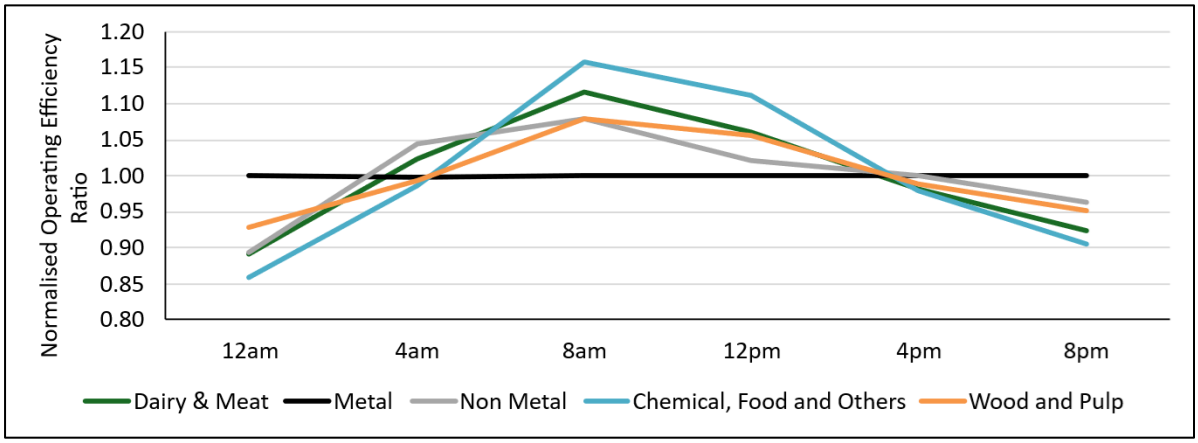


Figure 4.4: Normalised generalised operating efficiency ratio, prepared by the author using data from Forschungsgesellschaft für Energiewirtschaft (FfE) (2021).

4.4.3 Electricity Generation

Utility and Local Generation

The electricity generation sites are split into several categories, including electricity generated locally or utility-scale generation. For local generation, it is predominantly composed of residential and commercial solar generation, and it is consumed at a local distribution level. The approach is made so that the generated electricity will be consumed within each Point of Connection (POC), thereby removing the need for additional transformer setup as the generation is produced and consumed at the distribution level. At the utility level, electricity generation sites are usually operated by a gentailer, which serves as an electricity retailer and generation company or solely an electricity generation company.

Moreover, as for the local generation mainly residential and commercial solar generation, the generation at a node level can only be determined by population centres by calculating the net

outflow of electricity between each installation control point (ICP). Hence, the capacity build for solar in these population centres can be determined from the generation output.

As part of the estimation of the energy model network to identify significant electricity surplus areas in scenarios where the closure of hard-to-transition industries from fossil fuels is considered, consented and proposed electricity generation sites will be used as a potential location for green methanol generation as a proxy for green renewable industries for the future. Therefore, these potential sites would have the option to either be connected off-grid for power-to-fuel technology or connected on-grid for electricity transmission to demand centres.

As the model accounts for the cost of installed generation capacity as well as the operation of existing assets, the cost breakdown for generation is disaggregated into two components, the marginal levelised generation cost and new generation cost. Table 4.3 summarises the cost breakdown corresponding to the technologies. Cost parameters are primarily derived from New Zealand data sources. However, where local data are unavailable, estimates referenced from Australia and America are adopted as follows.

Table 4.3: Levelised cost for new generation.

Type	**Cost (NZD MWh^{-1})**			**References**
	Min	Max	Selected	
Hydroelectric	34.0	144.5	89.25	(IEA, 2012; Huuki *et al.*, 2022)
Solar Utility (CF=20%)	49.3	156.4	102.9	(Aurecon, 2024; Lazard, 2024)
Solar Local (CF=20%)	158.1	419.9	289	(Aurecon, 2024; Lazard, 2024)
Wind Onshore (CF=55%)	45.9	124.1	85	(Aurecon, 2024; Lazard, 2024)
Geothermal	108.8	180.2	144.5	(Lazard, 2024)
Bioenergy	102.0	493.0	297.5	(Visser, Thopil and Brent, 2019)
Thermal (Retrofitted)	187.0	387.6	287.3	(Visser, Thopil and Brent, 2019; Li *et al.*, 2021)

Selected value represents the midpoint of the reported cost range.

Important key considerations include the capacity utilisation factor or system capacity factor, which is an indicator of how much of the installed capacity is being used on average. The ratio shows the spare headroom available to accommodate peak load demand and ensure system reliability throughout the year. The ratio shows that the average demand and peak demand over generation from 2010 to 2022 are approximately 0.5 and 0.75 respectively (MBIE, 2025b). Therefore, the ratio serves as a proxy indicating that further electricity generation development is required, despite the existence of spare headroom, once the electricity system reaches its capacity utilisation limit. The capacity utilisation factor will be set as a global constraint in the energy model, and the P-graph representation as a modular operation within the generation component will be presented in the case study Chapter 6.

Intermittent Renewable Electricity Generation or Capacity Factor

Renewables.ninja is used to obtain intermittent generation time series for wind and solar. Because this step is part of the case study preprocessing, the full procedure is described later, including improvements made to automate data extraction for a list of generation sites rather than a single site. To streamline the workflows used throughout the thesis, the original source code written in R was converted into Python.

4.4.4 Biomass Resource

Biomass Types

Woody biomass, such as wood pellets, K log wood, and sawmill chips, has low moisture content and is comparable to coal briquettes and conventional charcoal. Hence, projects on retrofitting coal boilers are getting popular in site-level decarbonisation when moving away from coal usage, which reduces the carbon footprint from coal boilers.

On the other hand, agricultural residues such as in-forest residues, straw and stover typically are low-grade biomass resources with low processing capability. These fuel types with varied moisture content are more suitable to be fed into a bubbling fluidised bed for heat generation. Compared to biomass boilers, which are subject to more stringent operational demands, bubbling fluidised beds are more suitable for flexible operations with different variations in the heating value of the fuel. As biomass resource availability and cost differ across regions, the cost is not explicitly stated here but is detailed in Section B. Table 4.4 shows the solid biomass feedstock potentially used for thermal heating.

Table 4.4: Classification of boiler types by biomass feedstock.

Heating Technology	Feedstock Type
(Woody) Biomass Boiler	K-grade log wood, sawmill chips, wood pellets
Bubbling Fluidised Bed	Straw and stover, forest Residues

4.4.5 Electricity by POC-level

Grid Connection and Security Modelling

POCs in the transmission layer can be defined into four types: TEE connections, high-voltage direct current (HVDC) stations, grid exit points (GXPs) and grid injection points (GIPs). Fundamentally, these points define how electricity flows into, out of, and within the network. TEE connections are junctions where three transmission lines intersect, forming a three-terminal circuit. HVDC stations are electrical assets which connect to high-voltage direct current links to the alternating current transmission network. They usually function to transfer electricity and allow controlled power exchange between different parts of the grid, often transmitting electricity long-distance or connecting underground power lines. GIPs are substations where power stations supply surplus electricity to the transmission grid, whereas grid exit points are substations where electricity leaves the transmission grid to be distributed to local networks.

Each GXP acts as an aggregation hub for electricity delivery to local demand centres which are regional. The demand centres include electricity users such as residential, commercial and industrial consumers. As the energy systems model emphasises phasing out fossil resources on process heat in industries, there will be a trajectory where the electrification uptake in industries will increase drastically. Capacity upgrade is a fundamental consideration for the model. As the increase of deployment of electrical heating technologies such as heat pumps and electrode boilers, further grid asset upgrades are required from the transmission to the distribution level. In theory, N security and N-1 security are key reliability standards that are quantifiable, which ensure that the grid would be able to function without causing any outages in the electrical system despite a failure in n number of critical component(s). In short, N security refers to the condition where all grid components are functioning normally. Under this condition, the electricity network is operating at normal rated capacity. N-1 security is defined as a contingency situation where the grid is still operating and remains stable after one of the assets such as a feeder line or transformer malfunctions.

However, the increase in major grid components in the distribution and transmission levels to cater to high electricity use will increase the investment costs incurred for upgrades. In the model, elements such as transformer capacity upgrades, number of additional transmission grids required, new distribution lines in the factory level, construction of sub-transmission grids for new generation sites which are selected to operate on grids, are considered in the investment costs.

Grid Injection Points (GIPs) are also equally important in determining overall electricity resilience of the electricity system. The generation-side network security follows a similar principle. However, N security generation is commonly applied. This is because generation capacity is almost exclusively overbuilt to ensure grid reliability where peak electricity demand is met. Therefore, by incorporating the GIP nodes and its electrical constraints into the model, grid assets regarding the generation level are also evaluated to ensure that there is redundancy at the GIP nodes and stability during contingency conditions. The full list of pricing for grid capacity upgrades is presented in Appendix D. The investment cost shown is a Class 5 cost estimate, adopted from the cost estimation given by Ergo Consulting (Ergo Consulting, 2024).

When N security or N-1 security is not reported in the database, it can be approximated from electricity flows at each POC. The spreadsheet records electricity supply and demand as energy over a period, which is converted to an average power requirement by dividing by the hours of operation, yielding capacity in MW. The capacity is used to parameterise connection assets in a simplified way. Based on the required MW and assumed connection level, the spreadsheet assigns indicative transformer sizes and counts and selects representative distribution and sub transmission line categories as a maximum infrastructure requirement under full process heat electrification. This is simplified, as these are the parameters which account for the cost of upgrades in the grid.

Overall, rows correspond to POCs, and columns contain inputs and derived specifications. These calculations are completed prior to model execution because they are exogenous inputs at the

POC layer, ensuring consistent network and reliability parameters while the core model focuses on endogenous decisions.

Electricity Estimations at a POC level

Once the base year has been established, projections for future electricity demand can be developed. Four key sources of electricity demand need to be considered separately: residential and commercial users, battery electric vehicles, and industry.

Each sector is calculated or estimated independently. Residential and commercial demand is typically assumed to be proportional to population. Battery electric vehicle demand is estimated based on projected increases in electric vehicle market share. Industrial demand is generally assumed to remain relatively constant over time, unless there is a change in the process pathway.

Although the fuel input requirements change when an industry adopts a different technology or fuel type, which alters the amount of energy consumed from each fuel source, the end-use thermal demand remains the same, with the condition that there is no expansion or contraction of the site. As a result, the electricity consumption in the base year is treated as constant. Any additional electricity use that arises from the electrification of process heating technologies is accounted for separately in the model.

Estimations on Residential, Commercial and BEV Demand

Projecting electricity demand is difficult, especially in a spatially explicit way. Instead of estimating demand based on region or territorial authority, residential and commercial demand is more appropriately approximated using the number of ICPs within each POC. This approach is suitable because both the total electricity consumption and the number of ICPs for each connection are recorded, providing reliable data from the energy database. The reference year used for this analysis is 2022.

Other considerations include how the amount of electricity use varies in different regions. In particular, the Northern region consumes less electricity for heating, whereas the electricity consumption per household increases with lower latitude. Figure 4.5 shows the average consumption by colour scale for different regions.

The number of populations with commercial demands is assumed to be correlated to the number of installed ICPs. The population connected to the POC is estimated by having the fraction of ICPs in a POC of a region multiplied by the number of populations in the region. The calculation has to be regionally based as different regions have different populations connected to ICPs per capita due to the urbanisation of the region. As shown in Figure 4.6, the number of ICPs connected to the Auckland region is 2.39 people per ICP, whereas in Hawke’s Bay, the population to ICP ratio is 1.70.

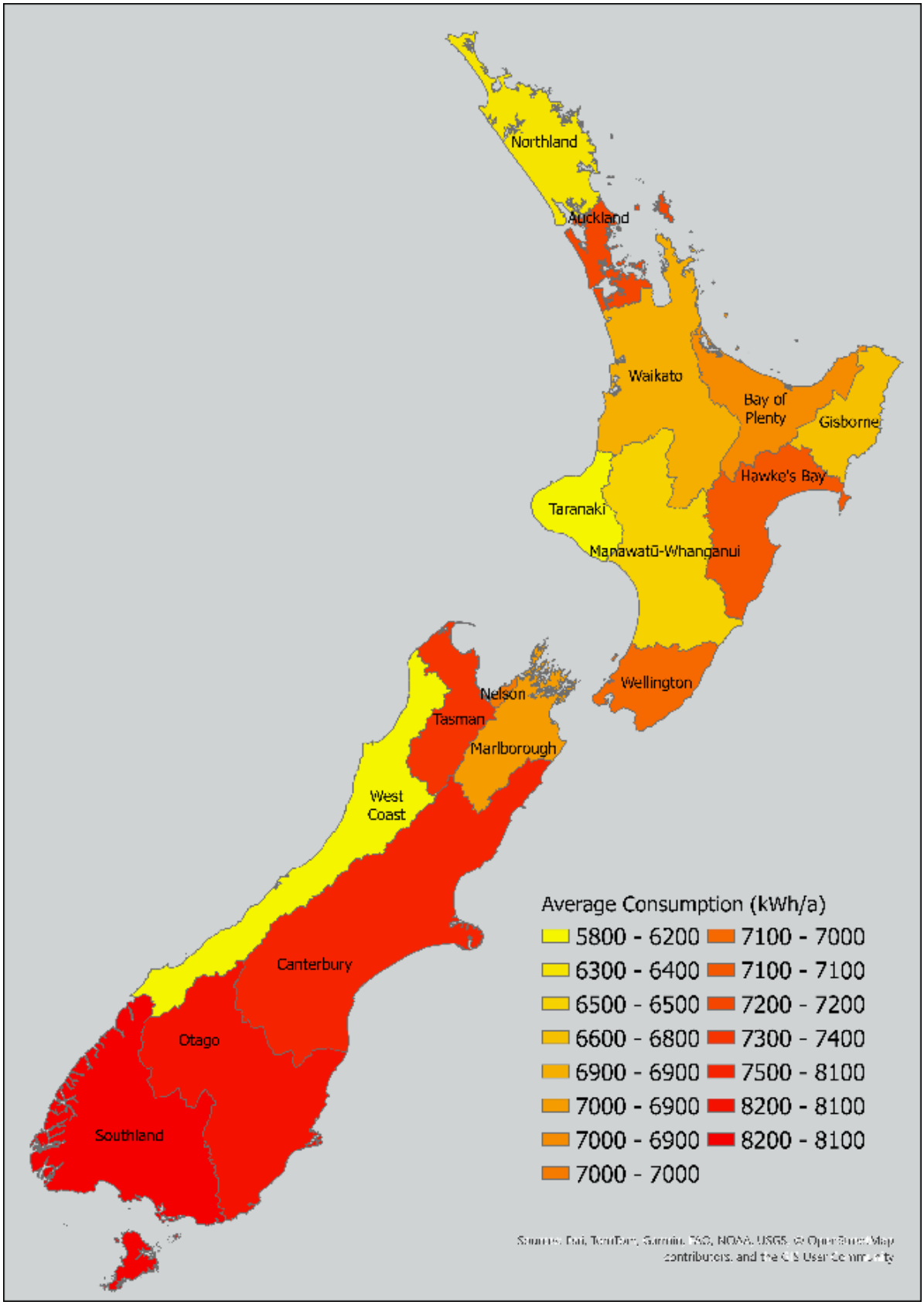


Figure 4.5: Average per capita electricity consumption by region, prepared by the author using data from (Electricity Authority, 2025).

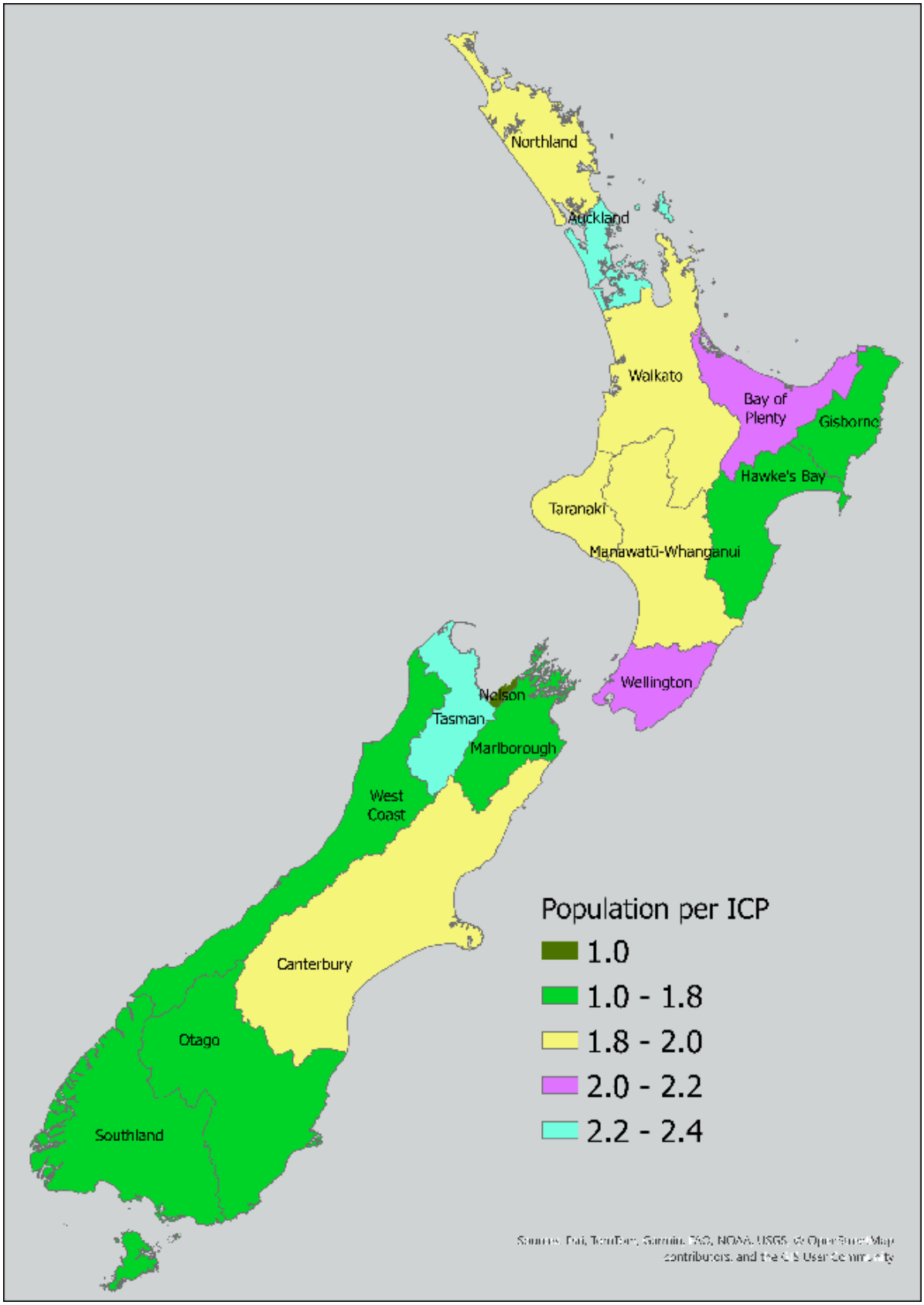


Figure 4.6: Average number of people per ICP, prepared by the author using data from (Electricity Authority, 2025).

After tackling the residential and commercial demand for each POC for the annual reference year, other features have to be attributed to ensure a good prompt for the energy model. These include the demand increase due to the rise of market share of BEV and the electricity consumption from the industrial side (Figure 4.7). These two elements had to be addressed in the reference year so that the projected growth of residential and commercial demand could be carried out by the population change in each POC, whereas industrial demand uptake remains consistent unless there is a major ramp-up or ramp-down in production, which might affect the overall energy use. On the other hand, the electricity consumption by BEV follows a different projection scale, as it follows a different growth rate. BEV demand is divided into each POC, as the former is proportional to the population number.

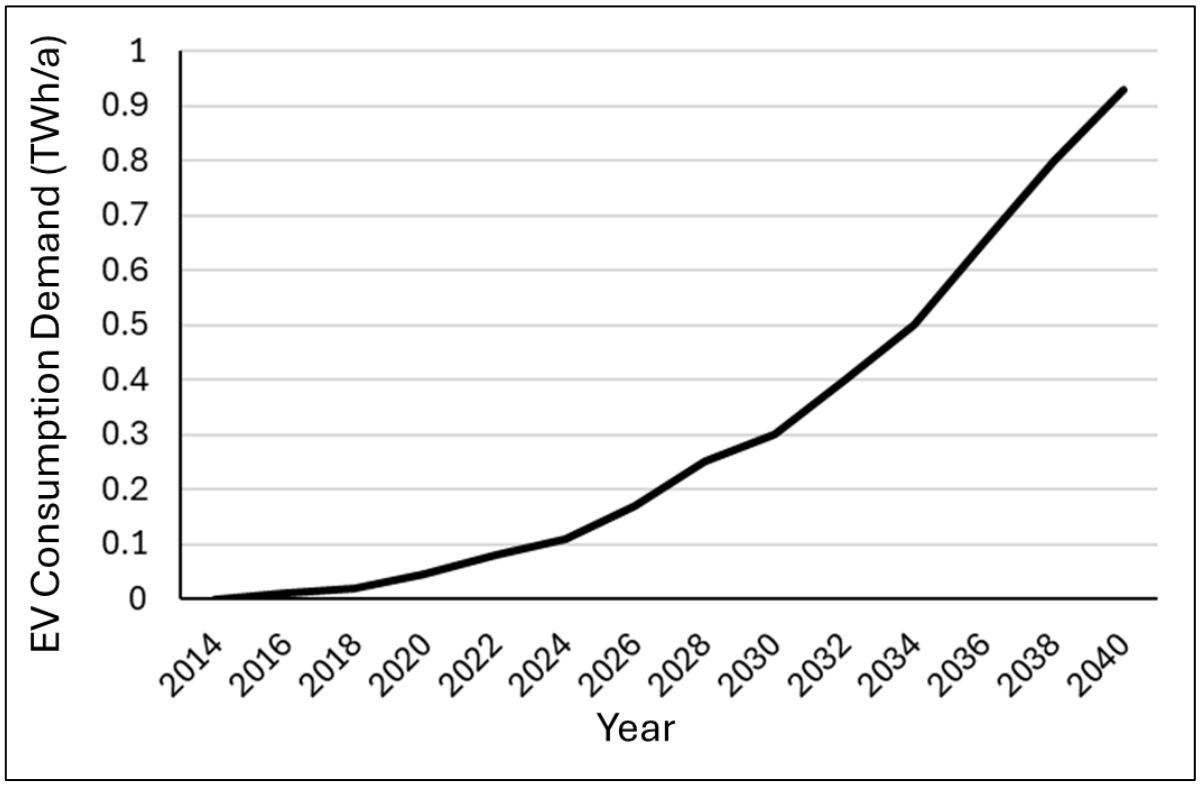


Figure 4.7: Projected EV electricity demand (TWh yr^{-1}), based on data from MBIE (2024b).

Baseline Electricity Demand in the Industrial Sector

As the historical total electricity demand for each POC can be retrieved from the Electricity Authority platform, and the electricity use for residential, commercial, and BEV sectors has been identified, the only remaining unknown is the current use of electricity in industries (Figure 4.8). Estimating electricity use in industries is complicated, as many uncertainties must be addressed. In essence, electricity consumption in industries differs depending on sector, process type, electrification uptake, the POC to which the facility is connected, and seasonal variations due to fluctuations in feedstock throughput.

To determine the factories connected to each POC, it is assumed that a factory is connected to the nearest POC unless specific information is disclosed. In such cases, the distance can be determined to estimate the remaining POC connections. The Near tool also returns the proximity

between two sets of points from different layers, which provides useful information for calculating the cost of line upgrades at the distribution level.

For residential and commercial demand, the monthly electricity consumption is estimated based on the proportional fraction of the annual residential and commercial electricity consumption. In addition, the annual electricity use for BEVs is distributed evenly across each month, as road usage is assumed to be relatively stable throughout the year (MBIE, 2024b).

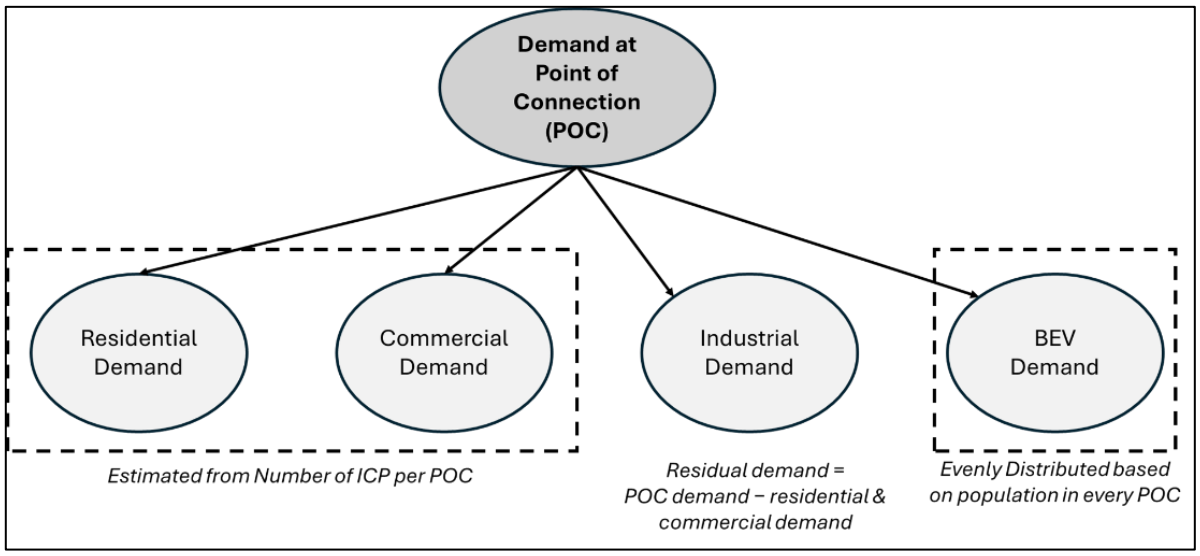


Figure 4.8: Conceptual representation of determining non-industrial process heat demand at each POC node.

In conclusion, the electricity demand estimation has to be conducted before employing the model as it is important to understand the amount of electricity to be consumed for the future projections, which tells us the grid allowance and necessary upgrades for implementing additional electrified heating technologies to decarbonise industrial process heat.

4.4.6 Green Industries/Methanol Production

Representation of green methanol in the energy model

To represent potential emerging green industries and competition for renewable resources, the model includes an optional green methanol production pathway. A sensitivity analysis is conducted on the selling price of green methanol to identify the price threshold at which green methanol production becomes economically viable within the national optimisation framework.

System boundary and siting assumptions

Green methanol production sites are represented as off-grid facilities co-located with electricity generation sites. This modelling choice is intended to reduce reliance on grid connection infrastructure and to avoid additional grid capital expenditure that could otherwise dominate investment costs for remote generation. Co-location also allows electrolysis to be supplied directly by on-site renewable electricity, avoiding the need for hydrogen transport infrastructure such as pipelines.

Within the optimisation structure, electricity generation sites provide the electricity input for hydrogen production via electrolysis, while biomass resources provide the carbon feedstock required for methanol synthesis. This coupling allows the model to evaluate spatially differentiated costs driven by biomass transport distances, local renewable generation potential, and competition for electricity supply.

Cost parameter handling

Cost assumptions are separated into global and site-specific components. Levelised installation capital costs and fixed operating costs for the methanol plant are treated as global inputs derived from literature values, while variable operating costs are site-specific. Variable costs include electricity input costs, biomass feedstock type and price, and biomass transport costs based on the spatial distance calculations used elsewhere in the model. This approach ensures that spatial heterogeneity is reflected where it matters most, while avoiding excessive parameter dimensionality.

Pratschner et al. (2023) provide reference estimates for green methanol production costs, reporting net present value values in the range of 569-785 EUR t^{-1} of methanol, which is equivalent to 1,100-1,530 NZD t^{-1}. Out of the cost compositions, it is estimated that the percentage for variable OPEX constitutes about 54.8% to 69.8% of the total levelised cost. These values are used as an external benchmark for interpreting model outcomes. For context, grey methanol is assumed to have a market price around 600 NZD t^{-1}, which corresponds to 30.2 NZD GJ^{-1} (LHV basis) or 108.5 NZD MWh^{-1} when using an LHV of 19.9 GJ t^{-1}. Under these conditions, green methanol is not treated as a cost competitive heating fuel in the base case. Therefore, within the model, green methanol is represented primarily as a product output rather than as a direct fuel input for process heat demand.

Process representation and conversion factors

The methanol pathway is not modelled using a full detailed reactor and unit operation flowsheet within the optimisation problem. Instead, it is represented using fixed conversion factors derived from literature and calculations (Pratschner *et al.*, 2023; Rajaee *et al.*, 2024). This approach allows methanol production to be integrated into the national system model while keeping computational complexity tractable.

For the cost of green methanol, the variable cost is calculated depending on the feedstocks (e.g. electricity and biomass resources) which are programmed in the energy model. Hence, the variable cost can be calculated by estimating the resources drawn from the electricity and biomass node. As a result, the cost function for the methanol component is assumed as the capital cost and fixed operating cost of the technology. Figure 4.9 shows the cost breakdown for green methanol production for a 50 MW_{el} plant capacity.

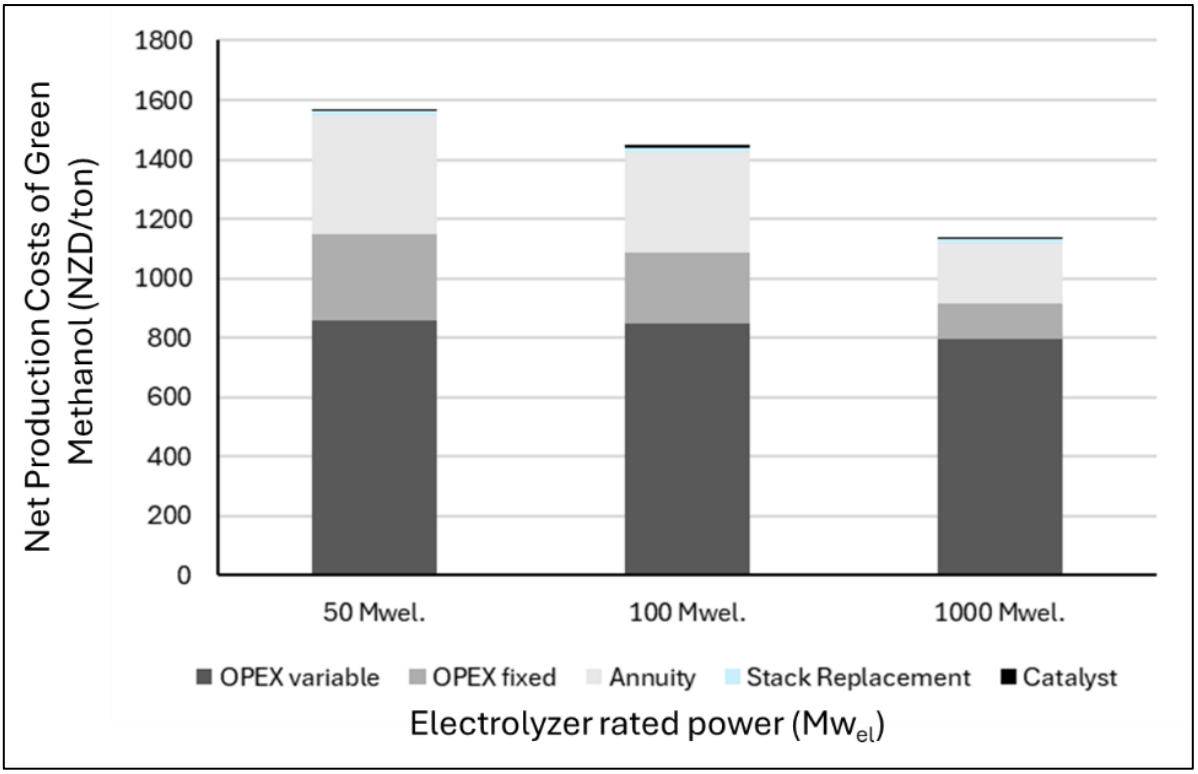


Figure 4.9: Net production costs of green methanol at different plant scales based on electrolyser-rated power, based on data from Pratschner et al. (2023), assuming 1 EUR equals 2 NZD.

The model uses the following basis assumptions:

- Lower heating value (LHV) of methanol: 19.9 GJ t^{-1}.
- Electricity input: 15.7 GJ t^{-1} of methanol produced.
- The electrolysis step accounts for 13.5 GJ t^{-1} of the electricity requirement.
- Biomass input is represented on an energy basis within the model, while methanol output is reported on a mass basis.
- The fixed capital expenditure (CAPEX) and OPEX for the power-to-liquids (PtL) technology are assumed to be 700 NZD t^{-1}(equivalent to 354.82 EUR t^{-1}), based on cost estimates for a representative 50 MW_{el} plant capacity (Pratschner *et al.*, 2023).

Because the system model represents biomass transport and electricity delivery using energy units, methanol is treated as a product that is tracked in tonnes for reporting and valuation. This process requires consistent conversion between mass and energy bases. The power equivalents are provided for interpretability and are derived from the same energy basis within the modelling horizon. The assumed energy and mass relations used in the model are summarised below.

Table 4.5: Conversion and energy basis assumptions for green methanol representation.

Basis	Mass (ton)	Energy (GJ)	Average power equivalent (MW)
Methanol product	1.00	19.90	5.60
Biomass input	1.42	14.25	3.96
Electricity input	–	15.70	4.36
Hydrogen input	0.11	13.53	3.76

Implementation in the optimisation framework

In the optimisation model, methanol production competes with other uses of electricity and biomass. Candidate methanol production nodes can be activated depending on their profitability under the assumed selling price and system constraints. This allows the model to identify regions where renewable generation surpluses and biomass availability make methanol production favourable and to quantify the opportunity cost of diverting electricity and biomass away from process heat decarbonisation.

Having defined the system layers and components, the next section explains the P-graph techniques used to represent them structurally.

4.5 P-graph Modelling Techniques

In this section, the P-graph bipartite graphical method is extended to represent multiperiod behaviour and cost structures, including both integer sizing and proportional capacity sizing. The novel contributions demonstrated in this thesis include capacity sizing with an explicit time dimension and operating unit selection based on integer unit specifications.

4.5.1 Time simulation model.

The time-dependent energy model is constructed according to the settings specified by the modeller. The time module decomposes and organises energy values and flows according to the chosen time resolution. This section provides a general explanation of the importance of time series modelling and the mechanisms through which temporal variation is represented in the system.

On the generation side, the underlying multiperiod capacity structure remains constant, while the actual electricity output varies with the capacity factor across seasons in an annual model, or with intermittency effects when an hourly time step is used. The installed capacities of technologies such as grid assets, industrial process heating technologies, and methanol production facilities are treated as time-invariant, enabling the optimisation to determine the optimal system-wide sizing.

The distinction between annual and time-series modelling is particularly important because residential and commercial electricity demand exhibits strong daily and monthly variation. Relying on annual averages would therefore lead to inaccurate system representation. Industrial demand, by contrast, is generally stable across the year as plants operate continuously to maintain efficiency, although certain sectors such as dairy and meat processing experience seasonal throughput variations.

To reduce computational effort when determining the operational schedule of available technologies at each time step, the model uses a synchronisation mechanism implemented through a joining node. This ensures that the capacity node remains active across all time periods, thereby preventing situations in which different technologies are alternately selected across periods in a way that inflates the combinatorial search space. This approach maintains temporal consistency while improving computational tractability.

The implementation for a time-step modelling using P-graph is presented in Figure 4.10. As an example, each arc connecting the intermediate node represents a multiperiod behaviour representing different time steps. The difference of energy from in this case can be represented by the coefficient or weightage of the arc. This time-step modelling structure can be further enhanced with the addition of capacity of technology for sizing and show better indicators of energy flow.

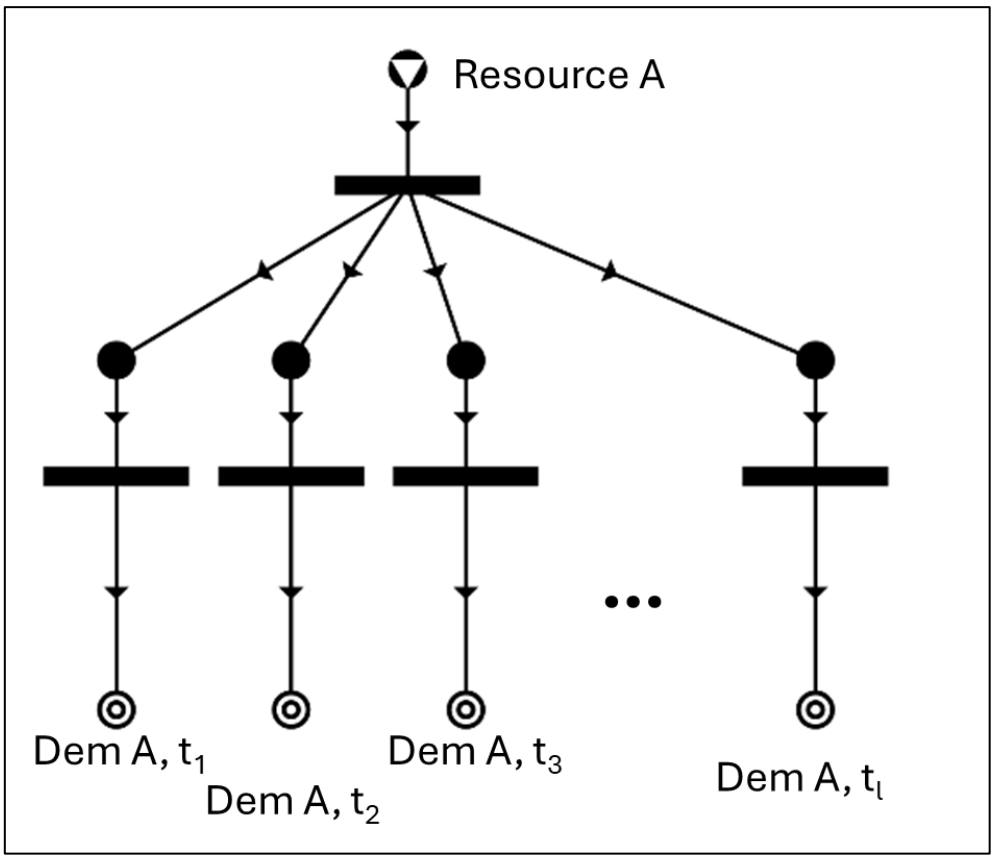


Figure 4.10: Illustration of time-step modelling applied in P-graph.

4.5.2 Proportional Capacity Sizing for Scalable Technologies

As an illustrative example, the time step properties incorporate sizing of technologies are displayed in Figure 4.11. As an example, the raw material node can either be a pool of resources all shared within the full horizon or could also be separate from the time-step showing that the resource is capped for every time step.

To ensure the flexibility of the capacity modular unit with time step element, it is designed to cope with varying energy flow across different time steps. Therefore, the energy flowing out of the capacity node into each time step has a coefficient multiplier to account for overbuilding to reflect realistic design for reliability considerations. The coefficient is also modified to account for the time resolution, as the batch of energy periods for sizing is different across different time steps. The intermediate node joined by the capacity operating unit has an upper limit to store unnecessary flows which are not required in the system without any penalty. In some instances, the individual operating unit across time steps can be associated with an operating cost whereas

the capacity operating unit can be accounted for by the capacity cost. However, in this energy model, operating unit cost is not accounted for here, as the generalised TAC cost is used in the capacity operating units, supported by resource nodes which represent variable cost based on consumption. This method was an improved method from (Tan and Aviso, 2016), where a recycle loop was added to reduce the number of subproblems so that the technology is selected in all time steps.

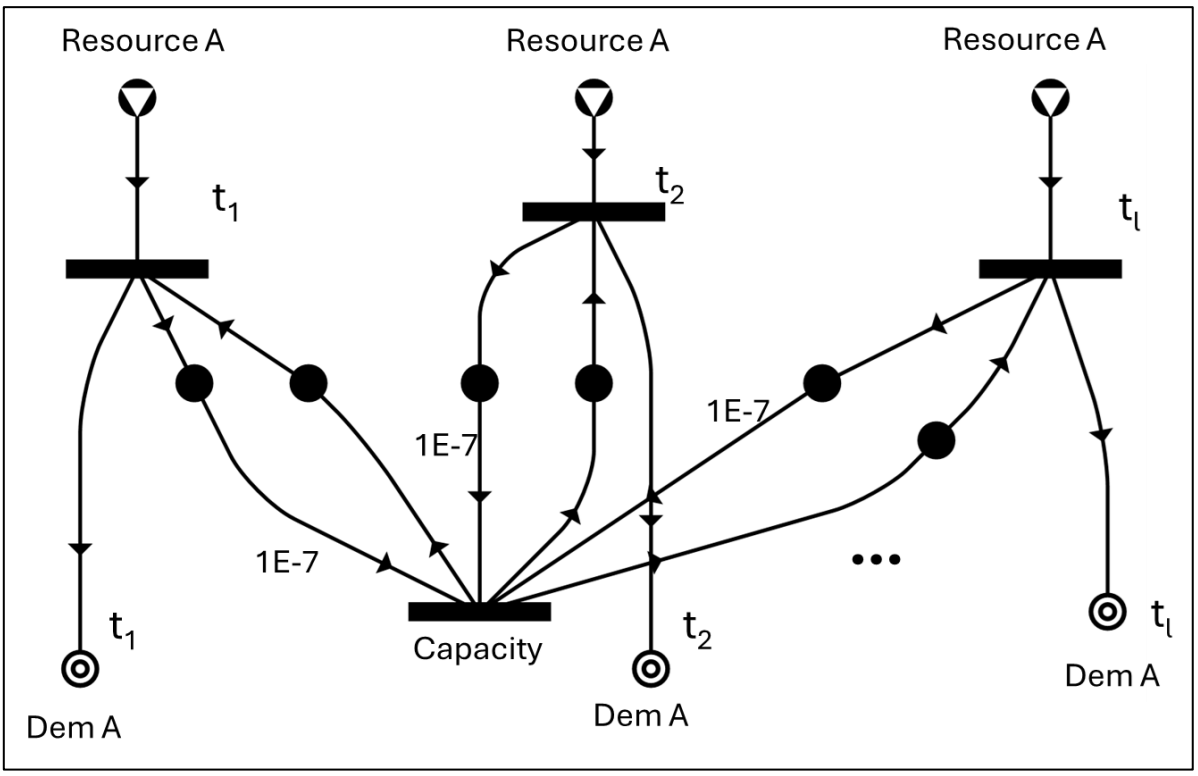


Figure 4.11: P-graph representation of a proportional capacity expansion problem.

In the context of P-graph shown in Figure 4.11, the "Capacity" operating unit node governs the energy flow by incurring a total annualised cost (TAC) cost by capacity sizing, associated with the largest energy flow required by the downstream demand. Hence, there is no restricted upper limit for the capacity, as it is used to account for the capacity needed to cater for the energy flow. An example is shown in Figure 4.12 on how the proportional capacity is modelled in that regard.

Based on the monthly time setting with time step $t = 3$, there is a 5 GWh per month allowance for time step $t = 1$ is 10 GWh per month for time step 2 and none for time step 3. In the real application, more considerations have to take place, such as oversizing the capacity of the technology for reliability and security of a system.

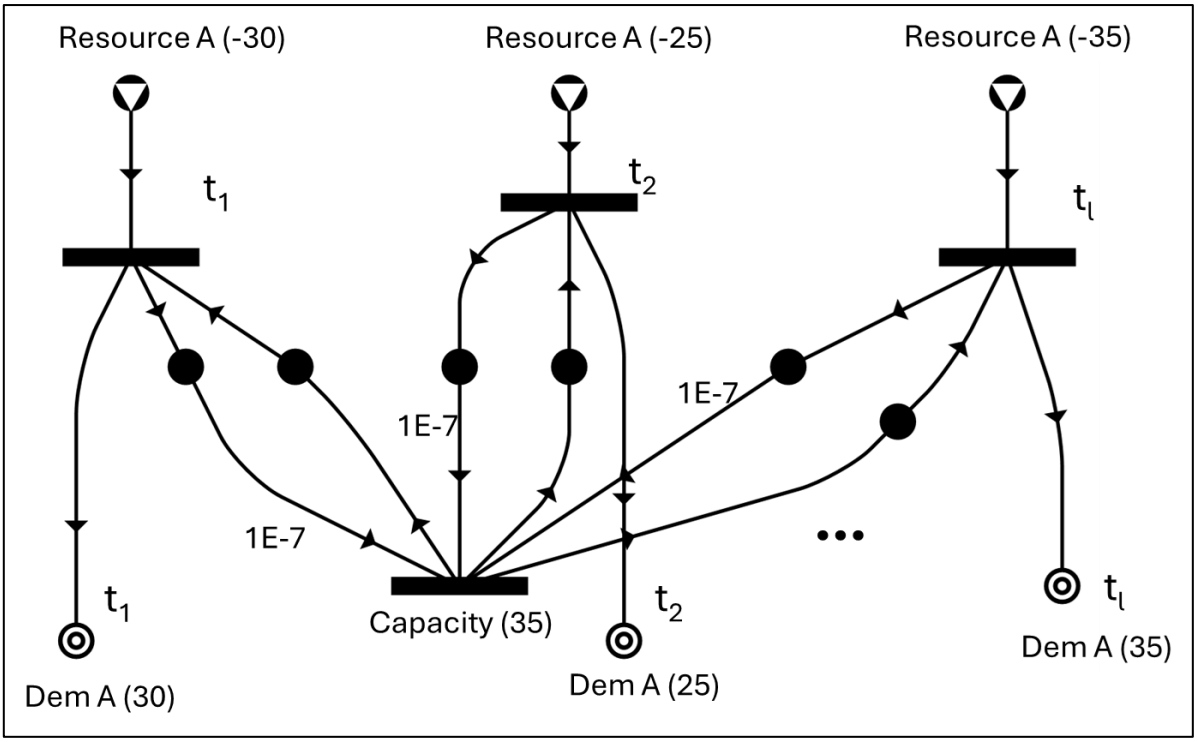


Figure 4.12: P-graph solution illustrating proportional capacity sizing.

4.5.3 Selection of Technologies Based on Integer Sizing

Sizing and calculating the optimal cost with a wide range of available technology proves to be a complex task. The sizing of a capacity of a technology does not solely rely on the consumption approach, as most technologies are not customised according to the required limitation. This is because technological assets such as transformers and distribution lines come in pre-defined sizings which serve different electricity flows and uses. In this energy model, a PNS approach is demonstrated to show how the technological selection is made as part of an infrastructural planning of the energy model.

This method of defining the types and number of technologies is similar to that used in the journal paper *Designing Optimal Small-Scale Solar Energy System by P-graphs*, which applies a two-level design planning approach to calculate the return on investment of a residential solar system. The study involves the optimal sizing of PV panels as well as inverters at a specific payback period, with the aim of maximising household profits through energy independence from the grid and the sale of excess electricity at a given rate.

A modification of the method in this application involves keeping the concept of integer sizing of technologies with a single-level approach, as the main technological barrier considered is the energy flow across a segment of a system. A novel approach includes considering existing capacity, where necessary upgrades are included in the model which incurs additional cost to the overall energy system. The sizing of electricity generation is applied only to proposed generation technologies that are present in the feasible solution. The approach with integer sizing will also consider the cost-effectiveness if a smaller-sized upgrade is necessary where only a few periods

require the larger-sized upgrade, which would be better to supplement the energy demand with an alternative pathway which is more flexible with change in requirement.

Figure 4.13 outlines the integer sizing approach for technology selection for a single time period. In the figure, one operating unit represents the allowable electricity flow from existing assets. A parallel operating unit is used to represent time step-specific electricity flow, where the connected material nodes must carry the same flow according to their fixed coefficients.

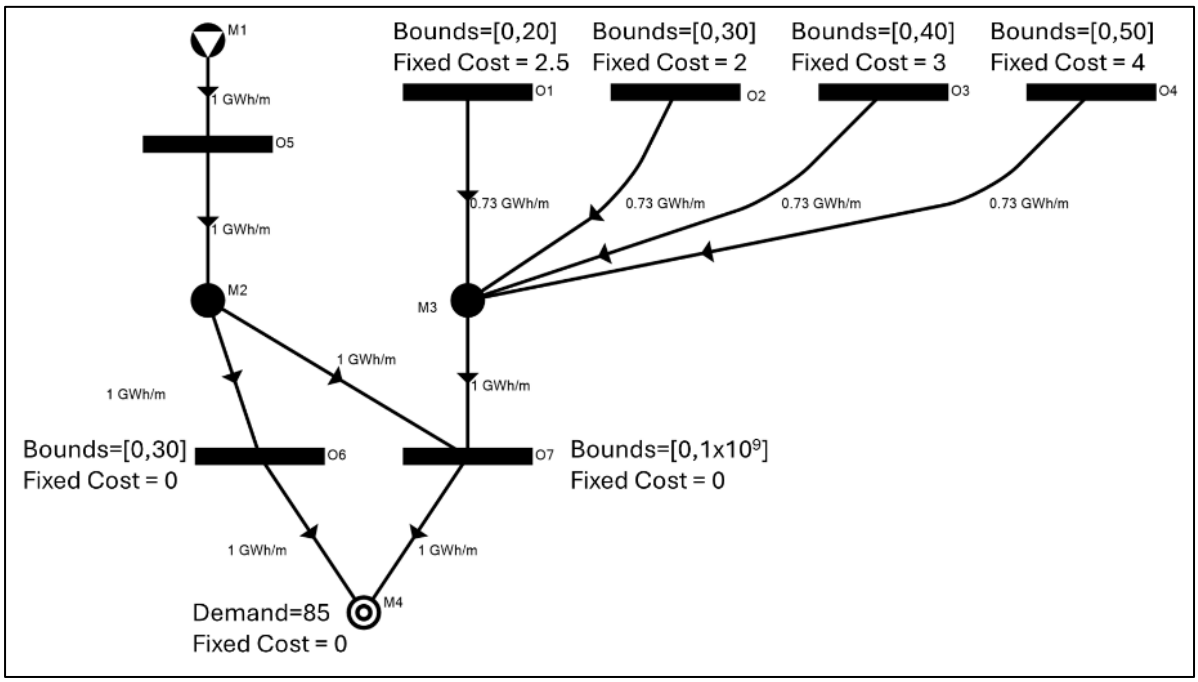


Figure 4.13: P-graph representation of an integer-based capacity expansion problem.

Operating units that represent potential capacity are also included. Each of these units corresponds to a specific technology and is associated with a cost function. Multiple operating units of the same technology can be used to represent the number of identical units installed, or the units can represent different technologies with different capacity limits. An inlet node is not required for operating units representing potential capacities, as this does not violate the P-graph axioms.

Figure 4.14 shows an example of integer sizing of a technology in the simulation, together with bounds and associated fixed cost. The potential technology capacity is in MW, with a time resolution of monthly hence the 0.73 coefficient for the edge shown in Figure 4.13. The demand is set as a value of 85 GWh per month in the example. Based on the Figure, the structure selected was an upgrade of 80 MW, this provides a relaxation of 3.4 GWh per month, shown in M3 node.

However, assuming the demand for $t = 2$ is 100 GWh per month, the structure will appear as a new configuration shown in Figure 4.15. From the interaction between time periods with varying demand levels, the optimisation model selects additional capacity to meet the higher demand, while spare capacity remains available during the first time period. By incorporating the interaction between integer capacity sizing and time periods, infrastructure upgrades can be estimated from the output of the energy model.

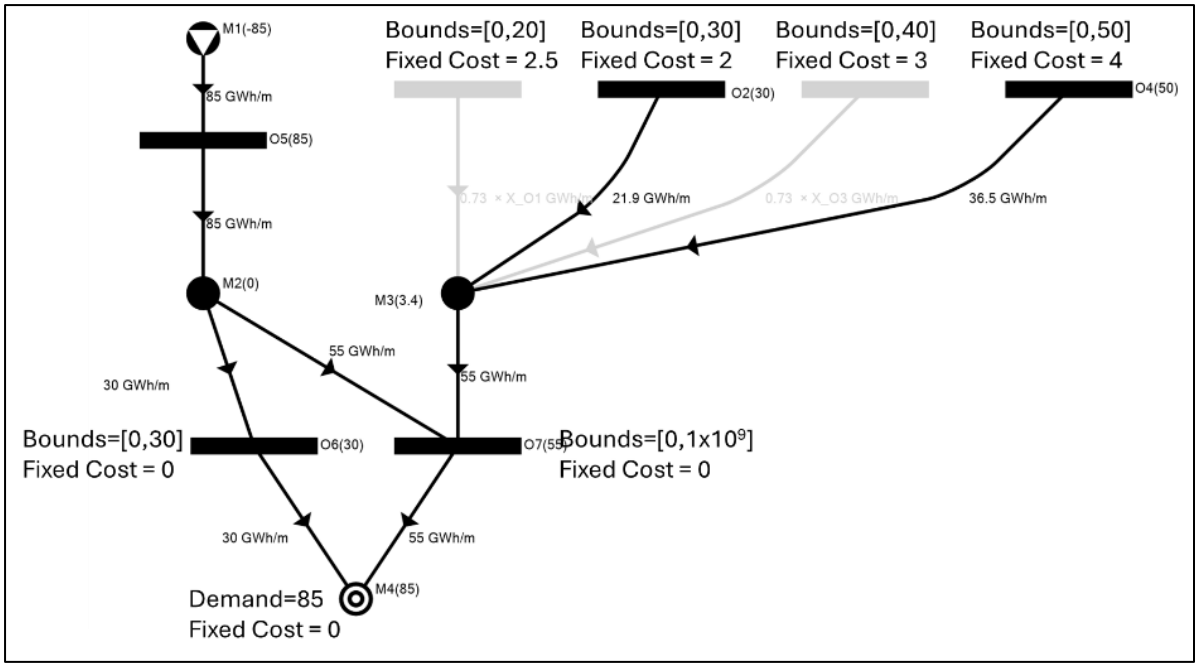


Figure 4.14: P-graph solution illustrating integer-based capacity sizing.

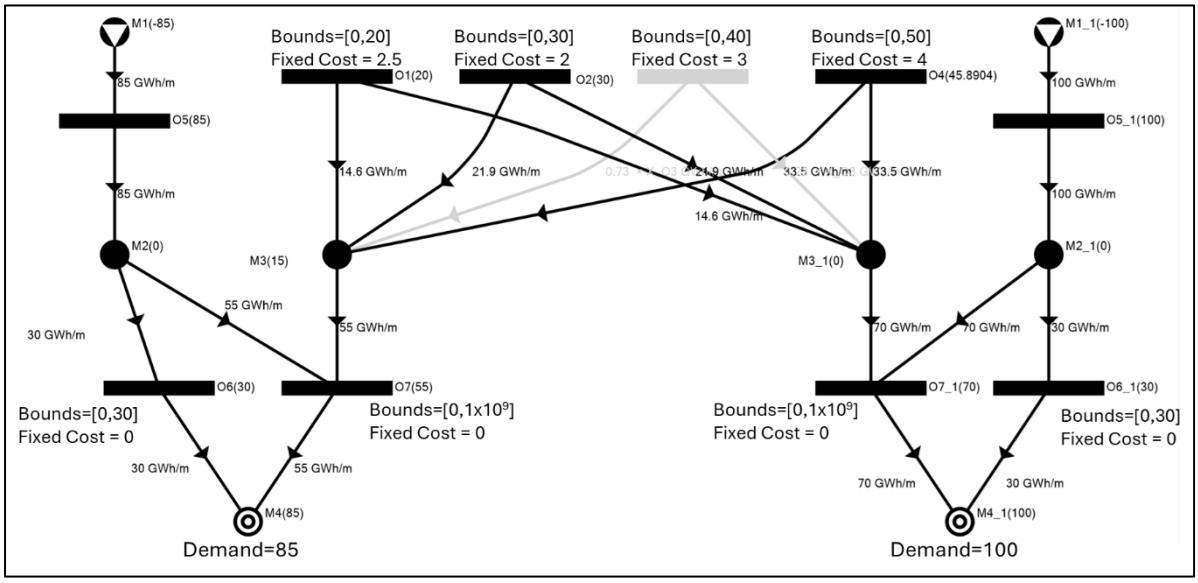


Figure 4.15: P-graph solution showing integer capacity expansion at time step $t = 2$.

The following section applies these representation principles to the reusable module templates required for the full energy model.

4.6 Modular Components Required for the Energy Model

In this section, the development of modular classes for the major operations in the model will be presented. Using Python, which follows an object-oriented programming paradigm, each class is designed to contain all essential information, including boundary constraints, flow connections between vertices, and the corresponding computational procedures. The modular operations described in this chapter include classes that represent heating technologies such as heat pumps, electric boilers, bubbling fluidised bed reactors, retrofitted coal boilers, and existing electricity grid components. The latter covers upgrades to grid assets such as transformers, distribution lines, transmission systems, sub transmission networks, and electricity generation facilities. In addition, the chapter will incorporate elements related to biomass transport for combustion-based power generation, as well as potential methanol production sites and the delivery of biomass to industrial plants. In this chapter, the P-graph framework for each modular operation is presented in detail to establish a clear understanding of the current modelling structure so as to allow readers to follow how the inclusion of additional iterations of similar operations in the large-scale model. The modular components available is summarised in Figure 4.16.

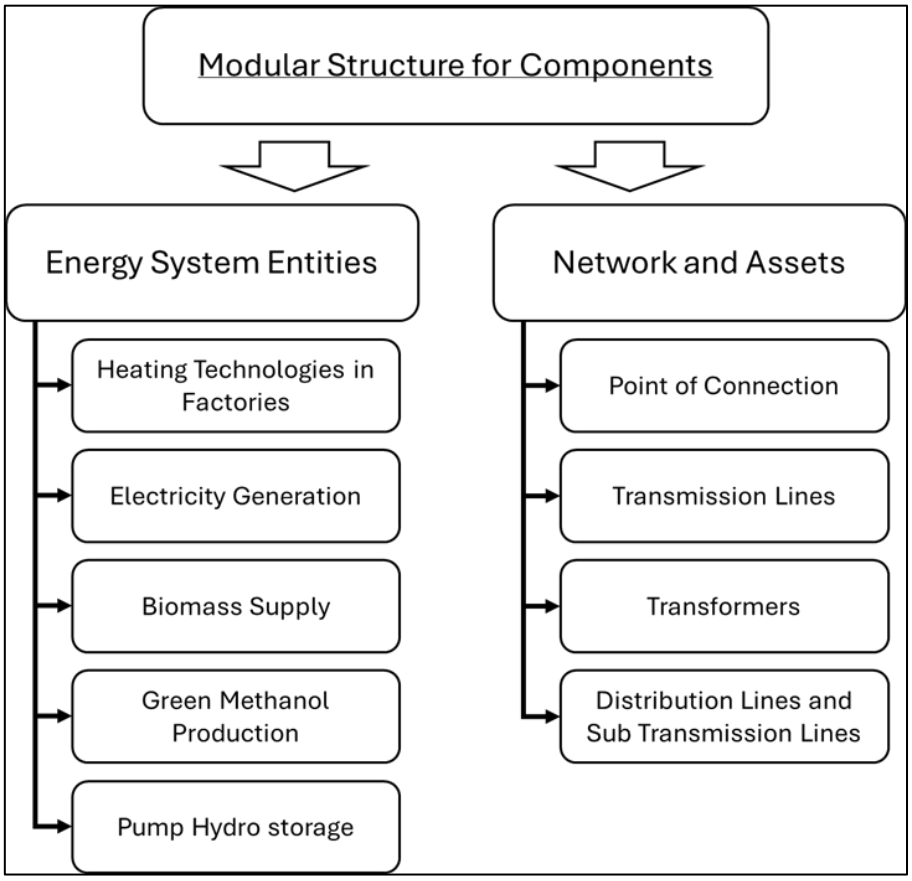


Figure 4.16: Modular components implemented in the energy modelling framework.

Example of Data Stored in Each Component

The study of heating technologies and input energy required for heating is one of the objectives to answer the sizing of heating technologies and fuel requirements to accommodate heating

demands based on their temperature levels. Figure 4.17 shows a simplified modular structure for process heat demands on the electrical side with important specifications included for process modelling optimisation but also for results presentation. In summary, the data stored for industrial process heat are site ID number, site name, coordinates, industry type, region, island, distance to POC, connected POC, total heat demand (GWh yr^{-1}) and demand at temperature levels. These stored values are important for specifying boundaries, and variables as well as vertices for each site are important to set up the problem table in the optimisation model. String values such as Regions, Island and Industry Type are stored so that these parameters can be merged with the optimal solutions values for each heating technology after the model has been solved. The attributes are important to evaluate the process heat flow based not on by site-level, but also from regional, island, and industry sector point of view.

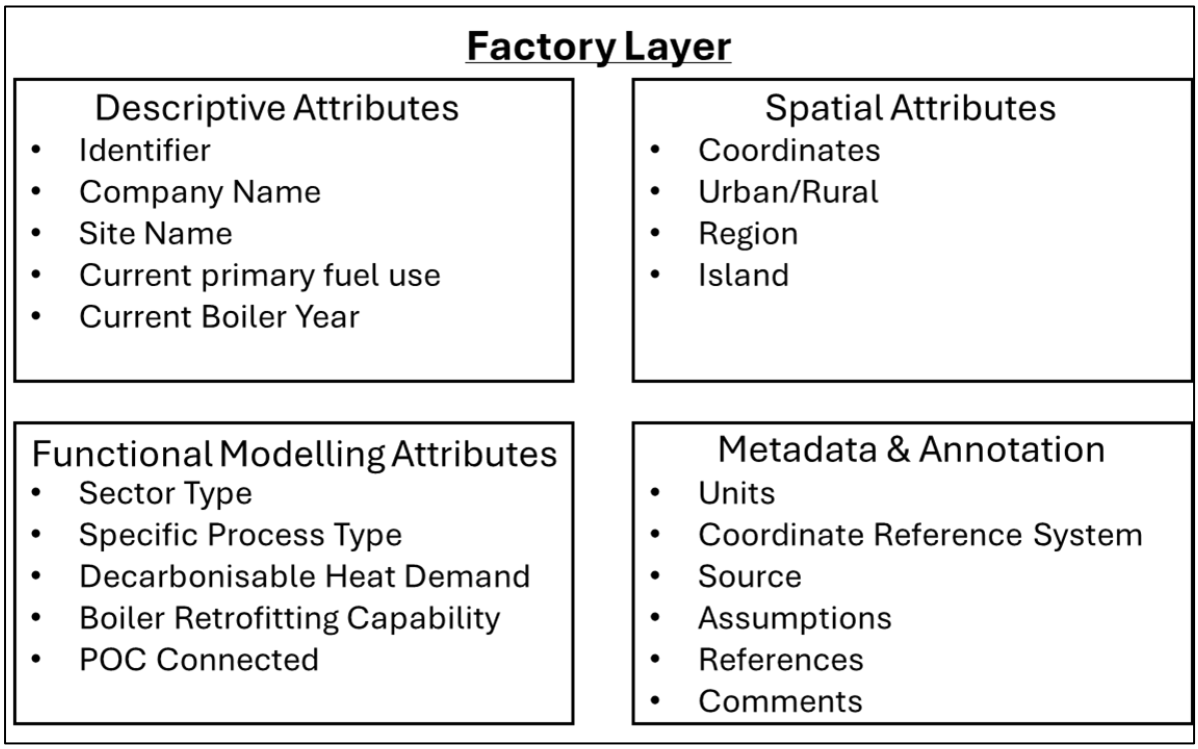


Figure 4.17: Example of component layer data structure, showing descriptive, spatial, and functional attributes, along with metadata and annotations.

In this chapter, the modular structure of the factory layer is outlined while the unit operations representing generation layer, biomass layer, hydro-storage, POCs and grid component are available in the Appendix F section.

4.6.1 Heating Technologies

The factory is connected by a unique POC, which transports electricity from the National Grid to the factory site. Each node is assigned with a unique identifier from a designated base value added to the node's index which is the Factory ID. The operating unit node labelled 103 is branched off to four different paths as for energy transformation change based on the different heating technologies corresponding to the temperature profile demands.

As the energy flow passes through node labelled 104+a, where a denotes the index of the heating temperature level, two outward edges can be seen branching out from the node. One branch connects to the node labelled 134+a, while the other connects to a set of nodes leading to nodes labelled 124+a. The coefficient on the edges 104+a show the efficiency and COP for their respective heating technology.

Product nodes labelled 124+a, represent the amount of decarbonisable heat demand in the model. For each of the product nodes, the lower and upper bounds are set to match the end-use thermal demand associated with the corresponding temperature level.

The operating unit nodes 134+a are representations for capacities build at different temperature levels of heat pumps and electrode boilers. Note that the operating units are joined by edges from different time periods, hence the optimal result returns the least-cost maximum capacity needed for the system model. The operating unit nodes representing capacities have an outward node with a weightage of a multiplier of the number of hours in the horizon with a reciprocal value for the coefficient of overbuild needed to represent actual design considerations when sizing the heating technologies.

As shown in Figure 4.18, for a monthly time step within a yearly horizon and an overbuild factor of 25%, the total coefficient leaving the operating unit nodes, 134 + a, would be 730 × 0.8 = 584, where 730 represents the number of operating hours and 0.8 is the reciprocal of the overbuild factor. Assuming a thermal efficiency of 1, this means that if the demand at a given temperature level is 0.730 GWh per month, the boiler capacity required for the site would be 1.25 MW.

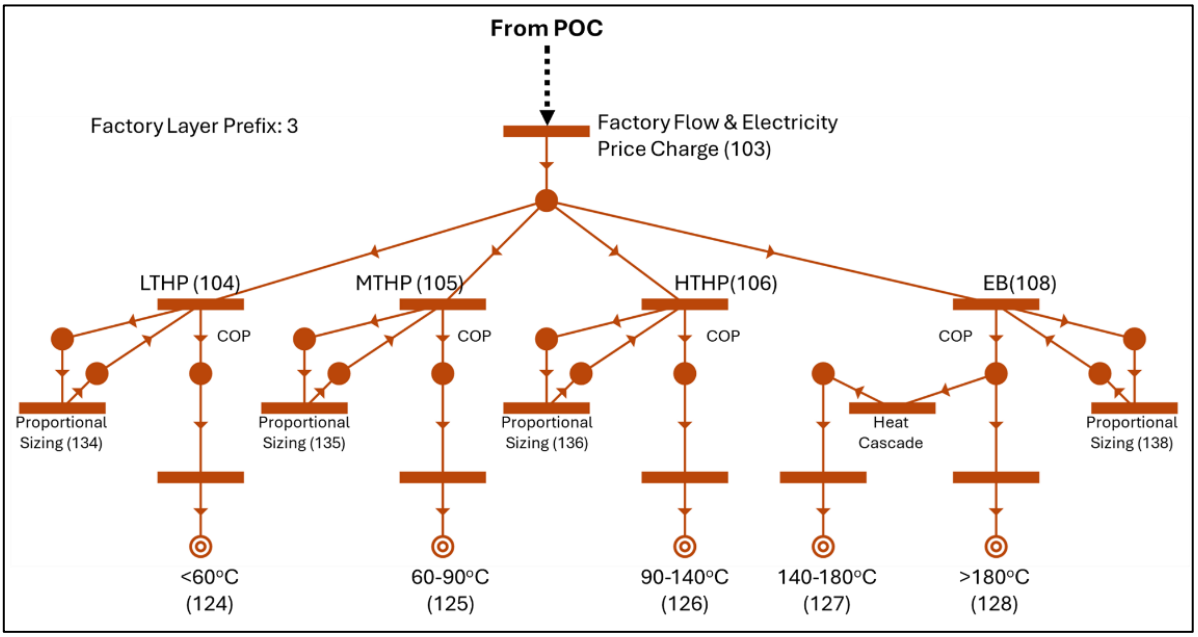


Figure 4.18: Modular component for the factory layer with |T|=1.

To reduce the computational burden of the optimisation problem, the heat cascade structure and the mutual exclusions associated with electrified heating technologies have been simplified. In the original formulation, heat flows from higher temperature levels could merge with biomass-derived heating at the highest temperature level. However, incorporating both mutual exclusions

and heat cascade constraints in the presence of biomass heat would substantially increase the number of discrete decision variables. These additional variables provide limited value within the current scope, given that the lower temperature demands are generally more thermally efficiently met by heat pumps.

For this reason, the model adopts an alternative representation in which heat supplied by electrification technologies and heat supplied by biomass technologies are directed through separating joining nodes. Although this approach increases the total overall number of nodes, it significantly reduces the number of feasible pathway combinations and therefore lowers the computational effort required to solve the model.

4.7 Chapter Summary

Chapter 4 has outlined the modular architectural design of the energy modelling framework, which establishes the premise of the modelling scope. The complex, multi-layered energy system model consists of several key components, including industrial process heat demand, electricity generation, biomass resources, and grid infrastructure, all within a single optimisation problem. A key contribution of this chapter is the definition of a superstructure-based modelling approach, where the function of each node within a modular component is explained in detail. The system configurations are represented to provide transparency and enable systematic exploration of fuel-switching strategies and infrastructure expansion options, which sets a baseline for the setup of the factory model in Chapter 7 and the grid reinforcement model in Chapter 8.

The chapter also introduced the concept of layered modular components, which represent either a technology block or a connection block between technologies. These components include factories, generation units, biomass supply, and transmission networks. All these components are flexible and reusable. Flexibility allows different types of technologies to be represented within a component through minor structural variations, while reusability enables iteration across large datasets.

In addition, the integration of spatial and temporal considerations within the modular architecture ensures that the framework can capture real-world constraints such as resource availability, technology sizing, distance-based logistics, and time-dependent demand variations. The representation of these elements within a unified structure addresses key limitations identified in existing energy system models, particularly the lack of spatial resolution in modelling industrial process heat at the site level.

Overall, this chapter establishes the structural backbone of the energy modelling framework, translating conceptual modelling requirements into a formalised architecture that supports large-scale system representation and optimisation.

Building on this chapter, Chapter 5 advances the framework into its mathematical and computational implementation. It introduces the formulation of the optimisation model, detailing how the modular components and superstructure defined in this chapter are translated into a solvable mathematical representation by integrating P-graph with Pyomo.

Chapter 5: Framework and Mathematical Formulation

5.1 Introduction

Chapter 4 established the modular classes representing the major operations of the model. This chapter focuses on the computational and mathematical implementation of the framework defined in Chapter 4. Building on that architecture, it presents the spatial energy modelling framework developed in this thesis by integrating P-graph superstructure generation, an automated translation tool, and a Pyomo.GDP optimisation model for the structural optimisation of large-scale energy systems.

As a continuation of the previous chapter, Python, which follows an object-oriented programming paradigm, is used to design classes that capture all essential information, including boundary constraints, flow connections between vertices, and the corresponding computational procedures. This is achieved through a porting tool that extracts P-graph specifications and converts them into an algebraic formulation in Pyomo. In this chapter, a novel approach to applying P-graph, differing from the conventional method, is discussed. The P-graph framework for each node and edge is presented in detail to establish a clear understanding of the generalised disjunctive behaviour of the model. This enables readers to follow how the porting process translates the P-graph structure into a mathematical optimisation model.

A central contribution of this work is the bridging of the P-graph representation framework with the Pyomo solving framework. Combining these two frameworks addresses two key gaps in current energy modelling practice: (1) providing a structured modelling interface for complex optimisation problems while (2) enabling access to a wide range of open-source and commercial solvers. This integration enhances flexibility in model formulation and improves robustness in the optimisation workflow and solver configuration. The chapter also extends the development of the model output structure and post-processing framework within the energy modelling framework, including spreadsheet-based outputs containing problem specifications and results, as well as masking techniques for extracting layers from identifiers for subsequent result presentation in Chapters 7, 8, and 9.

This chapter presents the computational workflow used to implement the energy modelling framework. It begins with how data is imported to represent modular components and model layers, followed by the application of the P-graph Python integration, the Pyomo formulation, the P-graph porting tool to establish a generalised framework, before concluding with a simple illustrative example for Section A. The framework developed here is designed to be understandable, straightforward to troubleshoot, reusable, reproducible, and adaptable to a wide range of applications. The procedures adopted in the workflow are explained and justified individually to clarify their purpose within the overall methodological design.

5.2 General Import of Data

The energy model uses a dedicated data-import routine to load, preprocess, transform, and organise all input datasets into structured data frames. The routine reads multiple Excel sheets from user-specified paths, which could encompass data sourced from different workbooks. The sheets are loaded and their contents are transformed into Python dictionaries. In addition to reading comma-separated values (CSV) and Excel spreadsheets (XLSX) formatted inputs, the data command file is also designed to load datasets stored in pickle (pkl) format, which are efficient in loading large datasets, for example, stochastic input values generated from random variates from a normal distribution for Monte Carlo optimisation. The column titles in the spreadsheets are cleaned before being transformed into keys in the dictionary to remove symbols, numbers and other characters which might interfere with the syntax and naming conventions in the code structure.

In the case of scenario modelling, the prompts for changing variables and parameters can be established in two different methods. Specifications such as biomass energy content and biomass resource availability are derived from spreadsheets and pickle files, whereas electricity prices, which is fixed-price variable-volume (FPVV) options and studied year, are modified by changing a coefficient or variable name directly within the Python script respectively. Figure 5.1 summarises over the entire energy model, where each module deployment and input and output data are included.

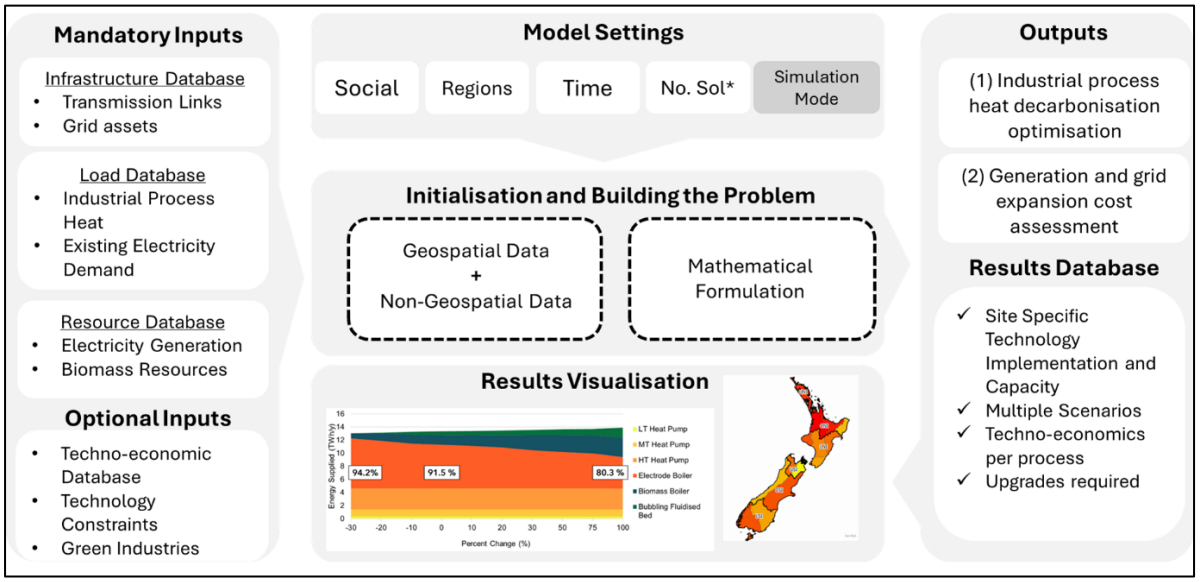


Figure 5.1: Model overview including required prompts, configurations and exemplary outputs.

5.3 Main Model Structure

After the data is imported, the command file applies a series of user-defined model settings that restrict the dataset to the relevant modelling scope. These filters may include temporal constraints, such as the number of time periods considered or the chosen time resolution, for example hourly or monthly values. The filters may also involve spatial criteria, including allowable transport distances or geographical boundary selections, as well as capacity-based thresholds.

These thresholds can include minimum demand requirements or minimum generation capacities that must be met before a unit is incorporated into the model.

From the model input point of view, scenario modelling is important because it allows the exploration of different futures and assumptions that influence the results when dealing with uncertainty in energy systems. Since this energy system model is fundamentally developed by a bottom-up approach for both demand and generation, it offers a realistic approximation of potential variation in energy requirements. The scenario set includes the electricity use for the base year and the future electricity use for the projected year. Residential and commercial electricity demand across all POCs is scaled based on population growth from the base year, while demand from industries remains unchanged.

The model also allows users to choose from a wide selection of biomass resource types, in which the flexibility enables a more critical assessment of the model, particularly when one or more biomass resources become constrained. For raw biomass options, availability is typically determined by the recoverability factor of forest areas, while wood pellets remain unaffected because they are export-based. Biomass supply scenarios are provided, representing a higher-limit and a lower-limit recoverability of forest resources. Biomass pricing is also a major decision factor when selecting pathways for process heat decarbonisation, and the prices for each biomass type can be adjusted for sensitivity and scenario analysis.

Other miscellaneous features of the model that are optional are the feasibility study on green methanol plants, produced by the power-to-fuel concept. The modular structure for the green methanol sites connected by biomass and electricity generation can be introduced by the user to explore locations with excess resources which could be leveraged to be a green methanol hub at a given selling price.

A unique identifier is required for each designated process modelling function to ensure consistent tracking throughout the optimisation workflow. Each index in every factory node, generation node, biomass node, and grid asset component is assigned with an identifier that is defined as a modelling variable to enable the solver to correctly interpret its bounds and specifications within the problem structure. These identifiers are also essential for post-processing, as they allow the model to extract and organise the relevant outputs before exporting the results for further analysis. The results are presented in an Excel format to support clear visual inspection and allow users to conduct data analysis. In parallel, key intermediate outputs are stored in a pickle format to enable efficient retrieval for the second stage of the optimisation workflow, which evaluates grid asset feasibility. The two-step structure allows the model to separate high-level system optimisation from the more detailed network assessment, improving both computational efficiency and analytical transparency.

In addition, the model gives the user the option to select from various available solvers, either existing P-graph-based solvers or commercial and publicly available solvers in Pyomo, due to the porting made between the two frameworks. These specification filters provide a clearer overview of the scope and context of the analysis and help manage trade-offs between computational performance and the quality of results.

Once model data are imported and structured, the next step is to connect the P-graph representation to a programmable workflow.

5.4 Coupling P-graph with Python

The application of P-graph, which was initially developed for solving chemical and energy conversion systems, has been extended to larger fields such as supply chains, waste management, resource allocation, discrete decision-making problems, hydrogen networks, evacuation route planning and many more. As the application and adoption of P-graph increase, the problem size has also expanded. The interface provided by P-graph Studio is no longer sensible for conventional manual data entry and preparation when dealing with large problems. A P-graph Python library has been developed which expands the problem size of the P-graph structure (Teng *et al.*, 2022). The P-graph Python library allows code interaction between the user and the library where an input file can be created and later fed to an executable P-graph solver of the user's choice. This approach allows coded users to generate nodes and arcs programmatically, enabling larger network structures to be constructed compared with the manual flowsheet construction approach. The workflow and usability of the P-graph Python library is displayed in Figure 5.2 and Figure 5.3.

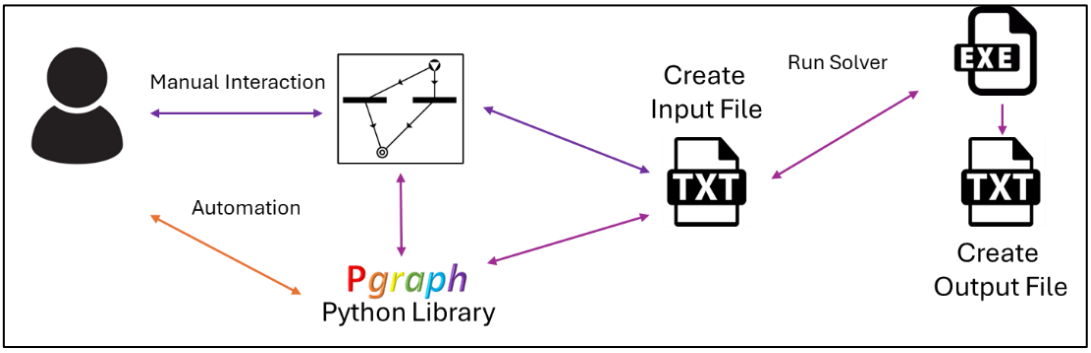


Figure 5.2: Adapted P-graph and Python workflow based on Teng et al. (2022).

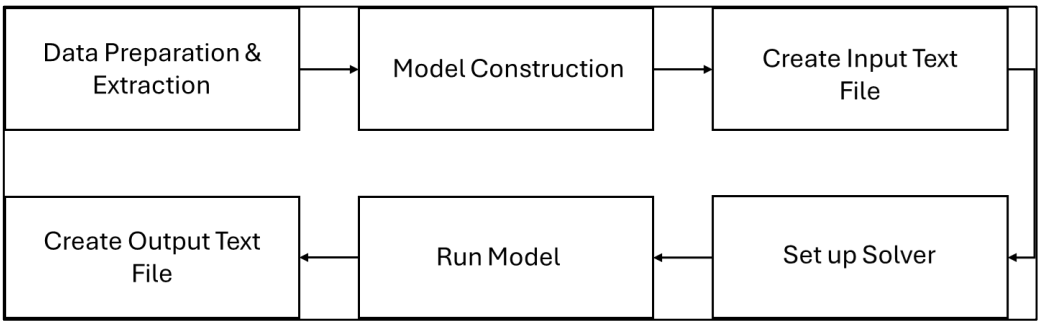


Figure 5.3: Standard P-graph workflow for process network synthesis.

Despite the significant progression of the new functionality in the P-graph space, the framework remains constrained by the P-graph executable where solver options are by default and cannot be modified by the user. The challenge becomes particularly significant when the solver deals with a large problem size with multiple discrete decision variables, which results in a long computational time. In such cases, the ability to define the optimality tolerance and the maximum

solver runtime becomes important to balance the compromise between the accuracy of the optimal solution and the computational effort. In addition, the executable is equipped with a 32-bit PNS solver, which has a solvable limit of approximately 200000 lines of syntax in the input file. As the model size increases, this limitation restricts the ability to solve very large systems using the native P-graph solver.

By leveraging the advantages of the PNS approach, the users would detect structural errors during model development in Python while supporting automation. A Python module was developed to convert a P-graph input file generated from the P-graph Python library to be ported into a Python-based Pyomo representation. The absence of the development of the P-graph to Pyomo port will result in users manually translating P-graph data into Pyomo syntax, which is both tedious and prone to errors.

The proposed approach in this thesis automates the translation process, enabling the optimisation problem to retain the structural benefits of the P-graph representation while accessing a wide range of commercial and open-source optimisation solvers. This extended workflow can be elucidated in Figure 5.4. Although the literature indicates that P-graph Studio provides access to alternative solvers, the generated equations have been found to be unstable, as they often fail to solve even small to medium-sized problems and offer limited flexibility in solver configuration. This limitation arises because the equations in P-graph Studio are formulated as mixed-integer linear programming (MILP) problems rather than in a GDP format, which can lead to computational inefficiencies.

In the next section, the development of the port and Pyomo optimisation modelling language, in particular, the General Disjunctive Programming approach is explained in detail as part of the development of the energy systems modelling tool.

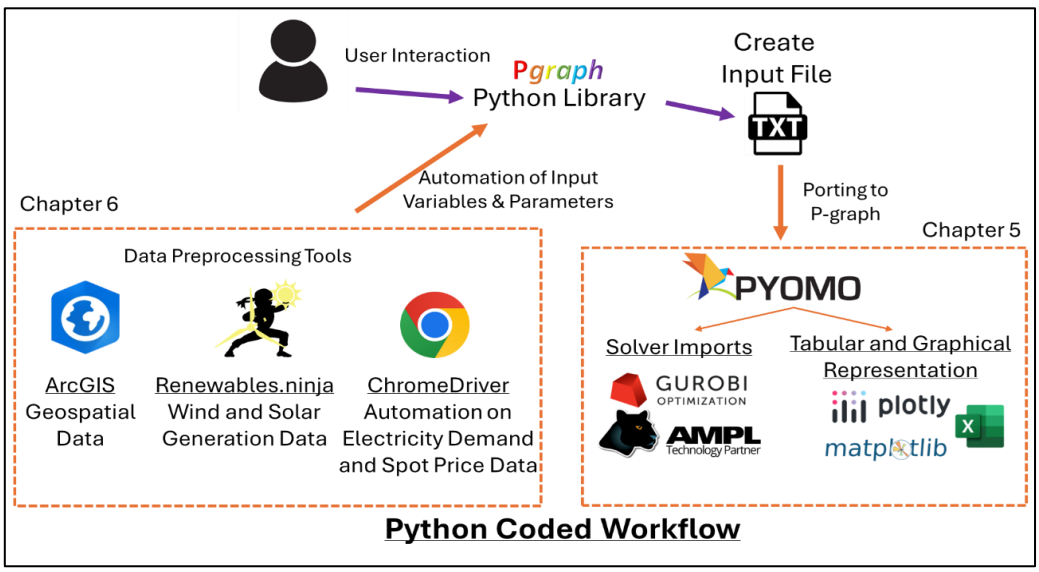


Figure 5.4: P-graph and Python workflow extended from the conventional version.

5.4.1 Application in the Energy Model

Before outlining the Pyomo formulation, it is useful to clarify the core idea of the approach. The framework first uses the PNS method to construct the energy system structure in P-graph. A porting tool is then used to translate the developed P-graph structure into an algebraic optimisation model in Pyomo, implemented using Pyomo's GDP framework. GDP is used to represent discrete technology and network decisions associated with the operating unit nodes and non-product material nodes.

The GDP model is subsequently transformed into a MILP problem by introducing binary variables and applying standard relaxation and reformulation techniques. This includes Big M and convex hull reformulations for the relaxation step. These methods are well established in the literature for transforming disjunctive constraints into a MILP form that can be compatible with commercial and open-source solvers. Once the formulation is reduced to a MILP problem, the energy model can be solved efficiently. This enables the optimisation to identify a feasible minimum-cost system configuration under the specified constraints. The simplified concept of the Pyomo backend formulation procedure is shown in Figure 5.5.

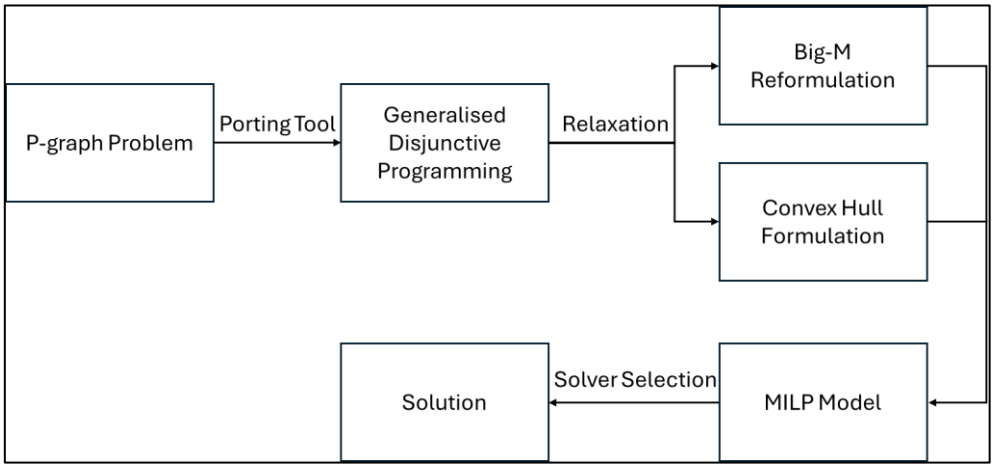


Figure 5.5: Problem formulation in Pyomo and the transformation process to a solvable MILP.

Disjoint rule between Operating Unit Nodes

The application of disjunction can be interpreted as mutual exclusion in the P-graph framework, where alternative options enforce the selection of one operating unit node within the process network. This disjunctive or mutually exclusive structure is important for selecting the most effective pathway in the network in accordance with the objective function. In the electricity delivery case, disjunction can be interpreted as enforcing unidirectional electricity transmission flows. It is enforced to reflect the physical behaviour of electricity networks where grid capacity constraints must be considered in the energy modelling of process heat industries.

On the other hand, disjunction or mutual exclusion can be applied in the selection of heating technologies in industry if necessary. This applies to enforcing mutual exclusion on heating technologies using the same fuel type but based on their respective operating conditions, on the temperature cascading operating units in each individual site. The addition of a disjunction rule

in these selections seems to be a useful feature to exclude two or more operating unit nodes from occurring at the same time for structural optimisation. As the energy model is set to be a MILP model, Equation (5.1) is used to represent linearity in the disjunctive model.

$$\bigvee \begin{bmatrix} Y_{ik} \\ M_{ik}x + N_{ik}z \leq e_{ik} \end{bmatrix} \quad \forall k \in K \tag{5.1}$$

Disjoint rule on Material Nodes and Operating Unit Nodes

In the combinatorially feasible solution network, the selected solution pathways are determined from the product node until they reach the resource node, once all constraints and conditions are satisfied. All lower and upper bounds of the material nodes must be specified before the solver is run to return an optimal solution. However, when material nodes are not included in the selected pathway, the solution value for the material node is 0, which is lower than the specified lower bound. Therefore, the disjunction rule is not only applied to evaluate technological selection within the energy model, but it is also important to allow relaxation of solution values in the mathematical system model. In particular, non-final product material nodes N, namely intermediate product nodes that act as preconditions of an operating unit along an edge that leads to one or more product nodes, apply a disjunctive rule at the node level. This allows the node to be inactive and discontinuous, so that it can take a valid value that is distinct and separate within a given range.

$$\bigvee_{i \in \{0,1\}} \begin{bmatrix} Y_{mi} \\ i = 0: \;\; x_m = 0 \\ i = 1: \;\; l_m \leq x_m \leq u_m \end{bmatrix} \quad \forall m \in N \tag{5.2}$$

Equation (5.3) is expressed when one or fewer than one branch is active:

$$Y_{m0} + Y_{m1} = 1 \quad \forall m \in M \tag{5.3}$$

Similarly, operating unit nodes that are within a set of edges, while adhering to Axiom 3, should apply the same formulation, as they are discontinuous within a specified range of values.

$$\bigvee_{i \in \{0,1\}} \begin{bmatrix} Y_{oi} \\ i = 0: \;\; q_o = 0 \\ i = 1: \;\; l_o \leq q_o \leq u_o \end{bmatrix} \quad \forall o \in O \tag{5.4}$$

5.4.2 Integration of P-graph into Pyomo

The novel contribution of this study is to bridge the gap between bipartite graphical optimisation representation and algebraic optimisation modelling by integrating the P-graph framework with Pyomo. This integration enables users to access a wide range of highly efficient commercial and open-source solvers while maintaining compatibility with the P-graph framework.

Another advantage of this integration is that it enables users who are familiar with graphical model specification to access the Pyomo environment indirectly. This is because both the conventional software-based specification and the Python-based implementation generate an

input text file that is subsequently passed to a solver, which then produces a text format output file.

A challenge faced by P-graph Python users is the readability of the raw output file. Consequently, data analysis cannot be conducted efficiently, particularly when dealing with large-scale optimisation models. The text file format is structured but difficult to interpret because the output values for operating units are embedded within flow equations rather than presented as separate entries. In addition, the output file is character-based, making it difficult to conceptually analyse the results from the optimal solution pathway.

One of the contributions of this work is the implementation of a structured Excel spreadsheet output format. The output spreadsheet also includes parameters such as Lower Bound, Upper Bound, and Price, which typically only exist in the conventional P-graph input file. Including these parameters in the output file provides an easier way to analyse the results. The improved output format enhances data accessibility and at the same time facilitates integration with external analysis tools used for data presentation and evaluation.

The generated output file consists of five main sheets representing material nodes, operating unit nodes and flow information, solution results and mutual exclusion data. Figure 5.6 displays the structure of the Excel Spreadsheet output format. In contrast to the conventional text-based output, the improved output format clearly separates variables, parameters and solver results into structured tables. The identifiers assigned to material and operating unit nodes may include prefixes, suffixes or infixes that denote their categorisation within a specific layer of the model. This structured naming convention allows the nodes to be filtered and grouped efficiently during analysis.



Figure 5.6: Comparison of output files generated by P-graph Python (left) and P-graph Studio (right).

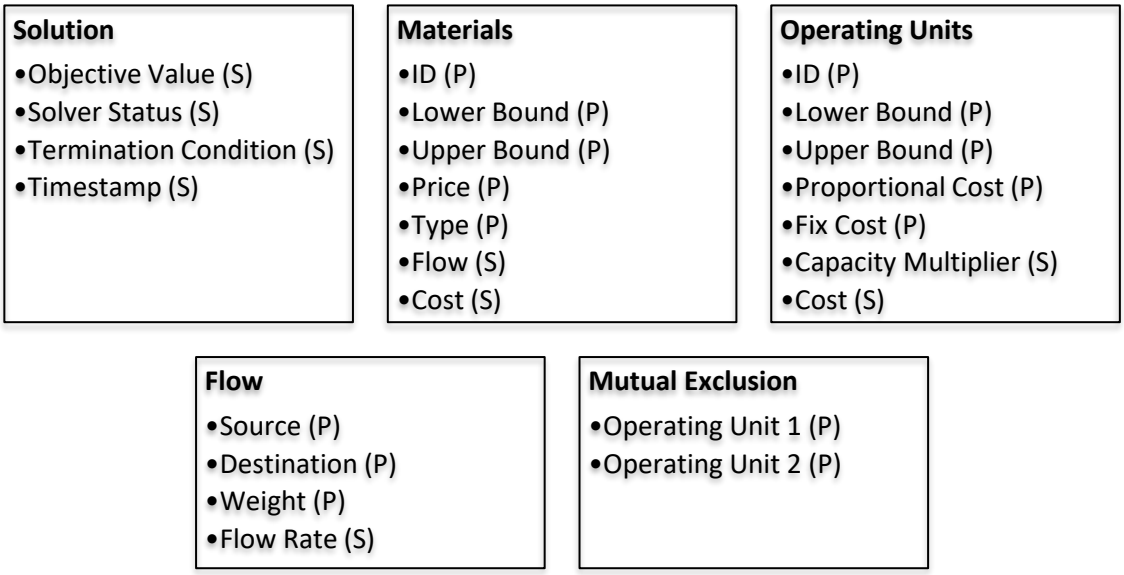


Figure 5.7: Implemented output sheets and their attributes in the Excel spreadsheet.

Referring to Figure 5.4 from Teng et al. (2022), the optimisation problem can be specified in two ways: (1) from P-graph Studio from a non-coded user or (2) from Python-based P-graph from a coded user. Based on the workflow by Python-based P-graph, this is followed by setting up the solver, which is constrained to four algorithms which return different outcomes.

The solver function is applied after defining the model to pass the input file generated earlier into an executable file. The executable file will return the optimal solution as well as a text-based output file for review and analysis. Moreover, the P-graph library could also generate the original network problem file, which is compatible with P-graph Studio for visualisation purposes. However, the P-graph Studio compatible file can only be opened for a small-sized problem network. The proposed workflow branches out once the text-based input file is generated from the Python library. The input file was then formatted and restructured to be compatible with Pyomo, allowing integration with the optimisation model. After establishing the coupled workflow, the translation mechanism is introduced to convert the P-graph specification into a Pyomo-compatible structure.

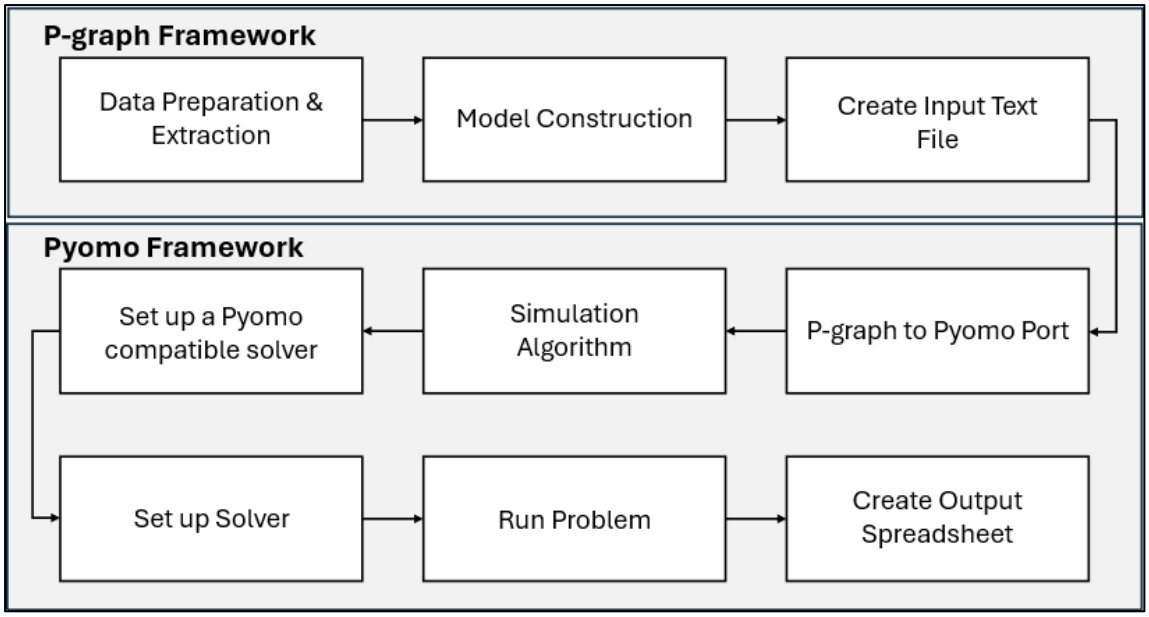


Figure 5.8: Proposed integrated P-graph and Pyomo workflow.

5.5 P-graph-to-Pyomo Port Tool

For any P-graph model, whether constructed in P-graph Studio or generated through the Python workflow, the problem is first written to a standard input file before it is passed to the P-graph solver. In this thesis, the input file is intercepted and routed through a port tool that converts the text-based P-graph specification into structured data that can be used directly in an algebraic optimisation model.

The porting tool reads the input file section by section, parses each line into its corresponding fields, converts the extracted values into the correct data types, and stores them in a dictionary-based representation. This is necessary because a single line in the input file may contain multiple specifications (for example bounds, costs, and types) that must be separated and assigned to the appropriate material nodes and operating unit nodes.

The primary function of the tool is to port a P-graph representation into a Pyomo formulation (Figure 5.9). In this workflow, P-graph provides the structural representation of the network, Pyomo provides the algebraic modelling interface, and the optimisation is performed using external solvers. Using the same principle in reverse, selected Pyomo results can also be exported into a P-graph compatible format to support scenario analysis and sequential optimisation (Figure 5.10), reducing configuration effort when repeated model runs are required. However, the reverse workflow shown in Figure 5.10 is not the main focus of the thesis. Nevertheless, the porting method is outlined in the Appendix A.

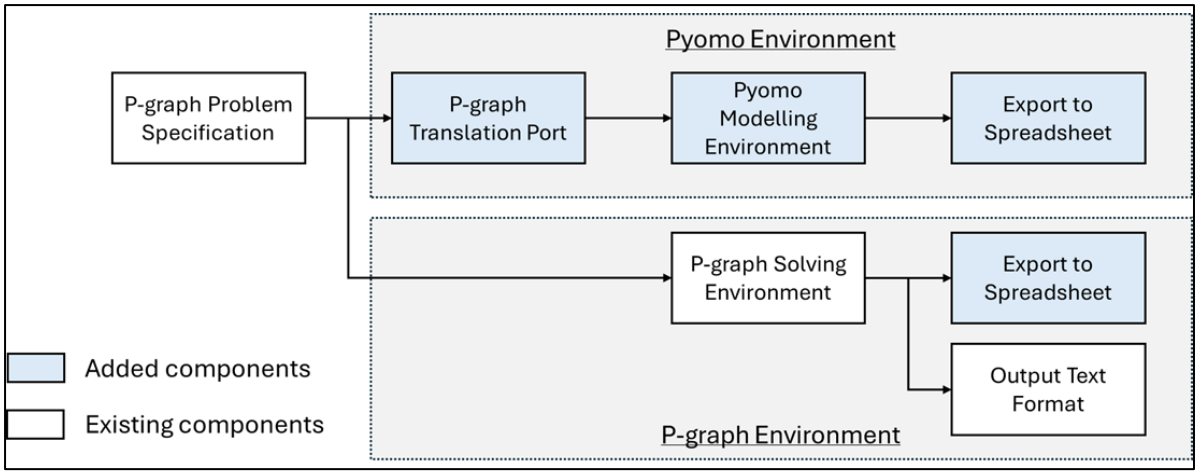


Figure 5.9: Options for modelling pipelines within the solving framework.

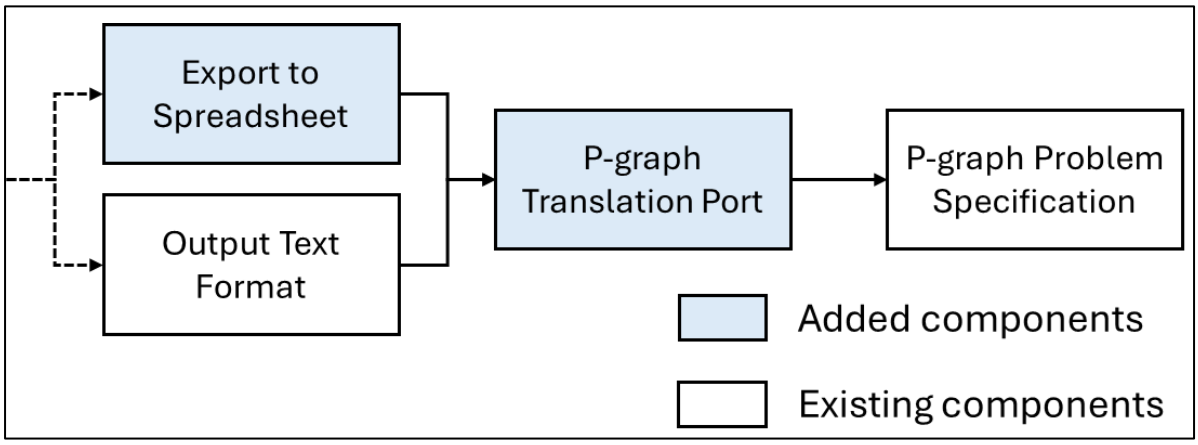


Figure 5.10: Modelling pipeline for the Pyomo-to-P-graph framework.

Overview for Switching a P-graph Problem into Pyomo Problem

The P-graph input file is generated after the problem has been specified and before it is passed to the solver. Input files produced by both P-graph Studio and the P-graph Python workflow follow a similar section order. The file begins with a header that specifies the problem name, followed by default units and default settings. The remainder of the file then declares material nodes, operating units, flow definitions, and finally the mutually exclusive operating unit sets (see Figure 5.11).

A key challenge is that P-graph Studio and the Python workflow do not use identifiers consistently. Although each node has an ID, P-graph Studio commonly writes the node name as the identifier in the file, whereas the Python workflow typically writes a structured node ID (for example, an alphabetic prefix followed by digits). Node IDs are preferable for automation because they are consistent and easier to handle when iterating many substructures within a superstructure model. Figure 5.11 illustrates the formatting differences between the two input file types. The porting tool is therefore designed to read and parse both formats before constructing the Pyomo representation.

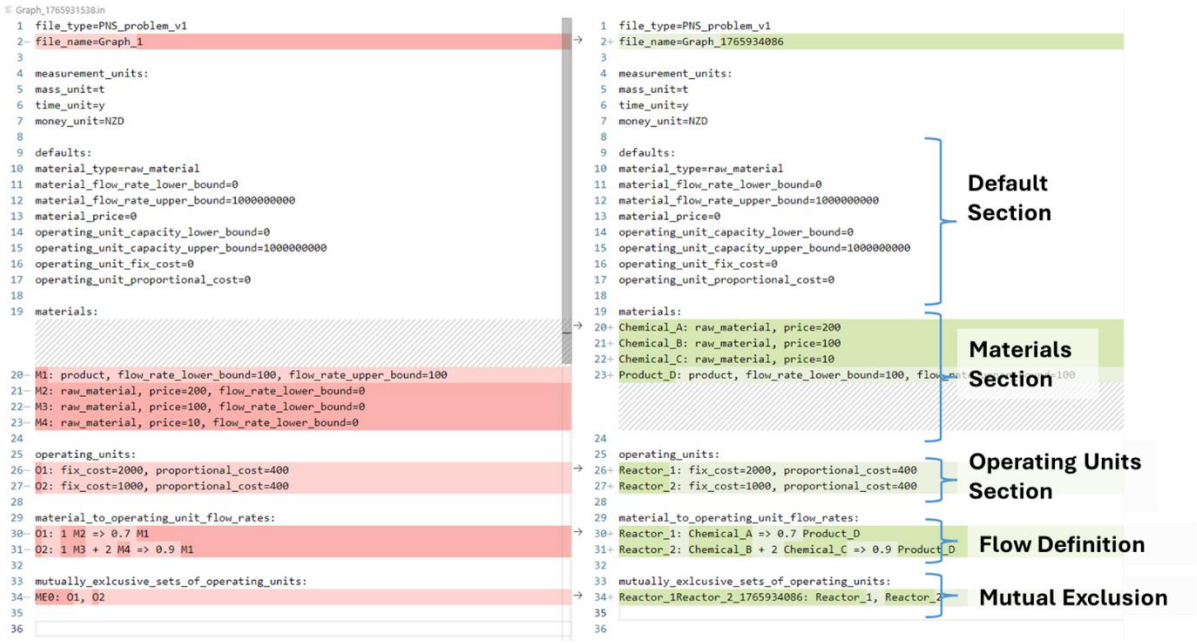


Figure 5.11: Comparison of input file formatting for P-graph Python (left) and P-graph Studio (right).

Regular expressions are used to parse the input files because both formats follow the same section order, but differ in spacing, naming conventions, and optional fields. To improve robustness, the parsing rules are designed so that formatting differences do not cause lines to be skipped or misclassified during parameter extraction.

Input parameterisation is split into five sections, following the order of the input file: (1) Default Section, (2) Materials Section, (3) Operating Units Section, (4) Flow Definition, and (5) Mutual Exclusion Section. The parser switches its regular expression rules when it detects the heading of each section. Only the key parsing rules are described here to explain the translation logic; implementation-level details are provided only where they are necessary to understand reproducibility.

Default Section

For the default settings section, a single format of regular expression is sufficient, as shown in Table 5.1. Default parameter names are formatted as alphanumeric strings with underscores that appear before the equal sign. The expression captures the parameter name as Group 1 and the assigned value as Group 2. The value pattern includes dots and sign characters to support decimal values and negative numbers. If the captured value is numeric, it is converted to a float.

Once the reading of lines reaches the materials section, a regular expression rule is enforced to capture all parameters of each material node from the script. The term where the rows matching the pattern are shown in Table 5.2 so that it can be easier to be explained and show examples which pass to it.

Table 5.1: Regular expression applied to parse default parameter assignments, including examples of captured groups and parsed results.

Lines	**re.compile(r'(\w+)=([\w\.\-]+)')**	
	Group 1: (\w+)	Group 2: ([\w\.\-]+)
material_type=raw_material	material_type	raw_material
material_flow_rate_lower_bound=0	material_flow_rate_lower_bound	0
material_flow_rate_upper_bound=1000000000	material_flow_rate_upper_bound	1000000000
material_price=0	material_price	0
operating_unit_capacity_lower_bound=0	operating_unit_capacity_lower_bound	0
operating_unit_capacity_upper_bound=1000000000	operating_unit_capacity_upper_bound	1000000000
operating_unit_fix_cost=0	operating_unit_fix_cost	0
operating_unit_proportional_cost=0	operating_unit_proportional_cost	0

Table 5.2: General regular expression rules applied to input and output files for data parameterisation.

Rule No.	**Regex Rule**	**Description**
1	(…)?	The symbol ? at the last term secures that the term is optional as the node will be treated as the default parameter specified in the first section.
2	\s*	\s* is applied for a good practice so that the number of spaces does not affect the match.
3	(?:,\s*)?	The term is an optional separated after a comma before the capturing the price. In most instances, the comma symbol is present for the specification if it is usually not positioned at the first group after the semicolon symbol.
4	?	The symbol ? is a quantifier, to make an occurrence optional.
5	(?: ...)	The term is used to create a non-capturing group, which groups things together without creating a backreference. This speeds up the grouping for each entry in this tule to group important parameters from the node.
6	-?\d+(?:\.\d+)?(?:[eE][-+]?\d+)?	The number expression is applied after the literal text to capture the number which could possess signed integers, decimals and scientific notation, where the -? sign is an optional minus sign, \d+ captures at least one digits, (?:\.\d+)? refers to optional decimal part with \. for decimals, \d+ for digits after decimals and ? to be optional, [eE][-+] captures exponent with signs whereas ?\d+ is accounts for the exponent digits. If \d* is present, it shows that at least zero or more digits can exist. The term is added in the flow section is identify to counter instances where no weight coefficient is before a material node which defaults it to be 1.
7	?P<…>	A general term used to indicate to Python that the matched item is grouped.

Table 5.2 (continued): General regular expression rules applied to input and output files for data parameterisation.

Rule No.	Regex Rule	Description
8	\w+	The rule allows the material ID to possess more than one or more character which includes alphabets, numbers and underscores.
9	text	The term is used for text which matches the exact literal text, ensures only the field name is matched and there is no ambiguity with other fields.
10	(?:_\w+)*	The term allows additional words including words starting with underscore to pass safely. The term safeguards material type like “raw_material” to be captured in the search process.
11	(.+)	The term matches the one side of the flow.
12	([\w]+)	Similar to term 8, but it creates a capturing group while keeping names or identifier of the text which includes a string of alphabets, numbers and underscores.

Table 5.3 and Table 5.4 list the general regular expression rules used in the parsing pipeline. The rules are designed to tolerate optional fields, variable spacing, and different numeric formats (integers, decimals, and scientific notation), while ensuring that each captured term can be safely mapped into the dictionary-based data structure used for model translation.

Table 5.3: General regular expression rule in the Material Section.

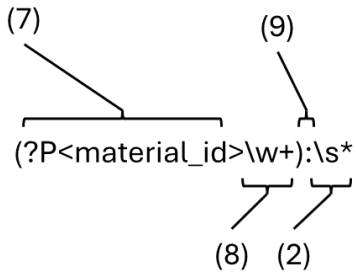
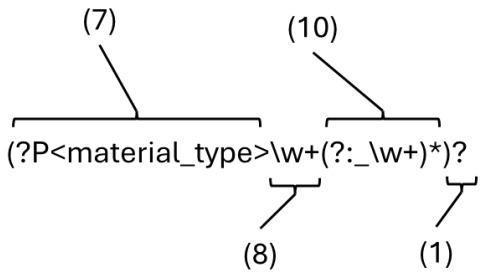
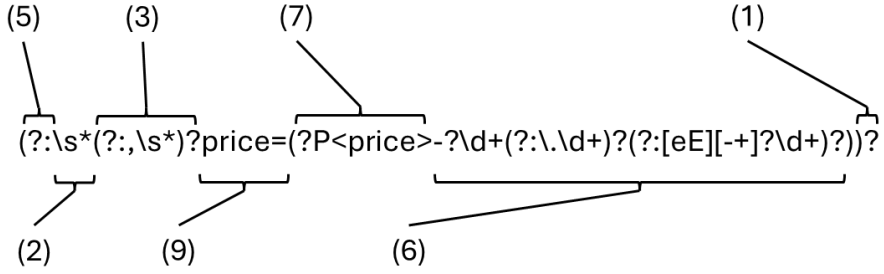
Specification	Structure
Material ID	
Material Type	
Price	

Table 5.3 (continued): General regular expression rule in the Material Section.

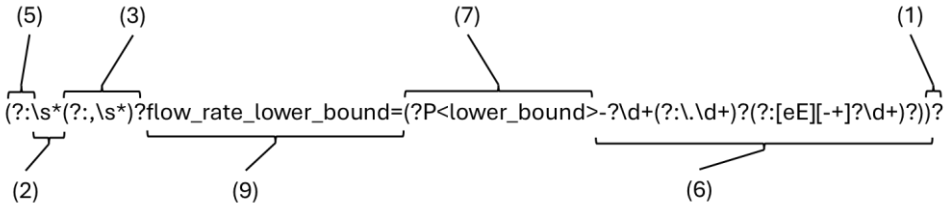
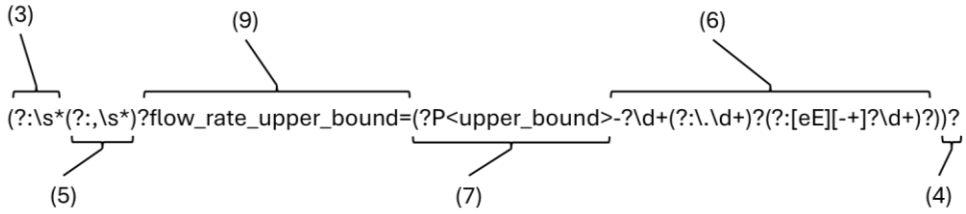
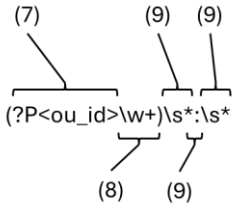
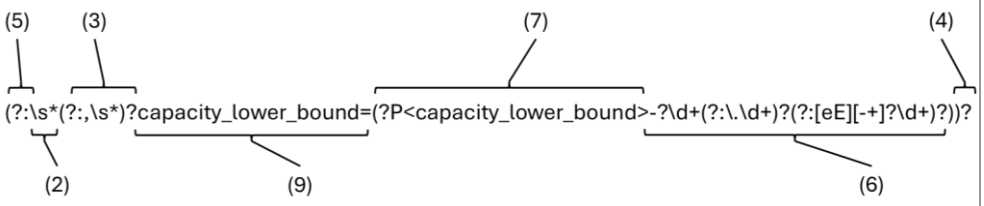
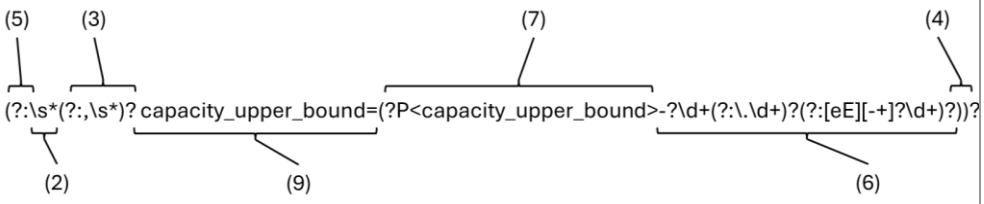
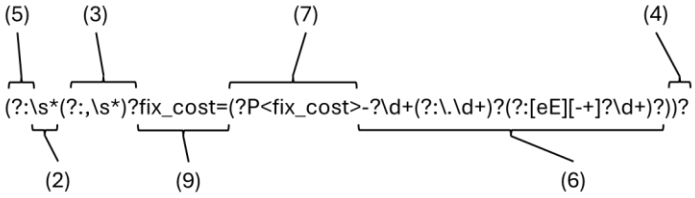
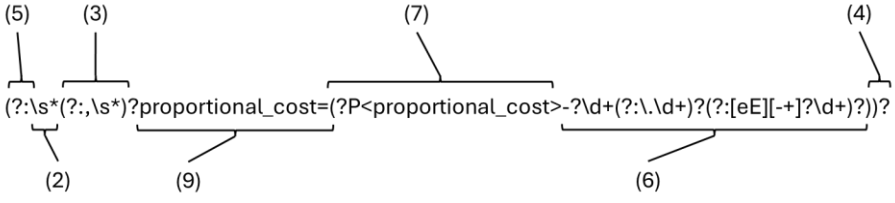
Flow rate lower bound	
Flow rate upper bound	

Table 5.4: General regular expression rule in the Operating Units Section.

Specification	Structure
Operating Unit ID	
Lower Bound	
Upper Bound	
Fix Cost	
Proportional Cost	

Flow Definition Section

In the input file, flow definitions can be difficult to parse because a single line may contain multiple inlet and outlet materials connected to the same operating unit (Figure 5.11). For stability reasons, the P-graph input format enforces a maximum line length of 16,384 characters. Lines longer than this limit may lead to unexpected behaviour or processing failures. Because the limit is character-based rather than node-based, a single line can still contain many material connections. This motivates a parsing approach that is robust to long lines, variable spacing, and repeated material terms.

In each flow line, the operating unit identifier appears before a semicolon. The input materials (with optional coefficients) appear before the arrow symbol =>, and the output materials appear after it. The parsing procedure therefore first splits each flow line into three parts: (1) the operating unit, (2) the input material list, and (3) the output material list.

After the input and output lists are isolated, each list is parsed using a regular expression that identifies an optional numeric coefficient followed by a material ID. If no coefficient is provided, the coefficient is set to 1.0 by default. Each parsed term is then stored as a directed flow tuple consisting of a source node, a destination node, and the corresponding arc coefficient. Materials that feed into an operating unit are recorded as arcs from the material node to the operating unit node. Materials produced by an operating unit are recorded as arcs from the operating unit node to the material node. Figure 5.12 shows an example applying the regular expression.

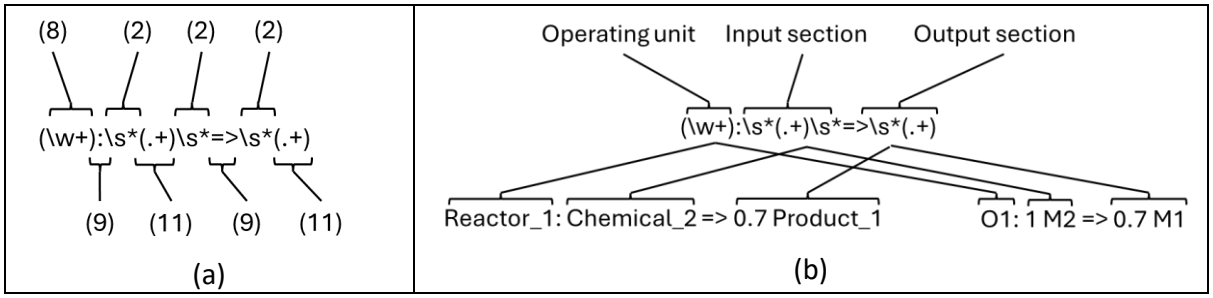


Figure 5.12: Regular expression rule enforced in the Flow Section, where (a) shows the general regular expression rule and (b) shows the example and grouping names.

Mutual Exclusion Section

For mutually exclusive sets, the disjunct index is recorded up to the semicolon, followed by a comma-separated list of operating units. Two regular expressions are used to remove optional whitespace so that spaces are not treated as part of the identifiers. The operating units are stored as a nested list in Python, which provides a consistent structure for passing mutually exclusive groups into the algebraic model. Figure 5.13 shows an example of a mutual exclusion general expression applied in the modelling framework.

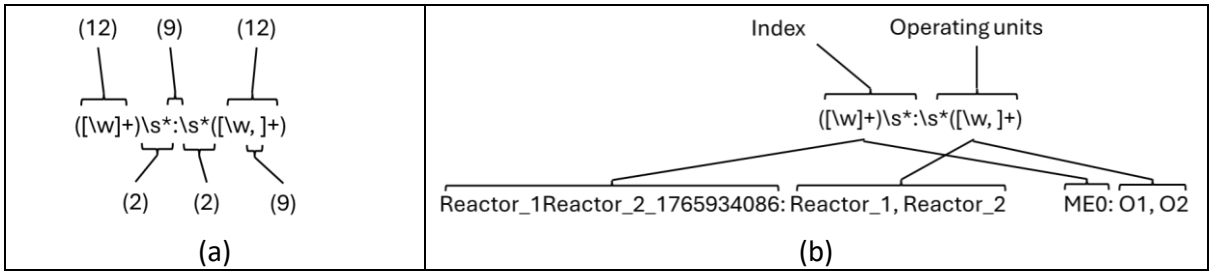


Figure 5.13: Regular expression rule enforced in the Mutual Exclusion Section, where (a) shows the general regular expression rule and (b) shows the example and grouping names.

5.6 Pyomo Model Structure

Once the input file has been parsed, the extracted values and their associated parameters are assigned to the corresponding Pyomo variables and parameters. The Pyomo equations were made to be compatible with the disjunctive P-graph framework. The Pyomo structure is built in Visual Studio Code and is available upon request. In the making of the energy model, the Pyomo structure is organised into five core components, where each component with its own conceptual role. They are categorised as (1) Sets and indices, (2) Parameters, (3) Decision Variables, (4) Constraints, and (5) Objective Function. In this section, the axioms and P-graph superstructural network will be mentioned corresponding to the components represented in the Pyomo framework. The Python code is provided in Appendix E.

Index Sets

The sets section defines the indices and categorises within the problem model. It consists of primary sets for materials M, non-product materials N, and operating units O.

$$m \in M \tag{5.5}$$

$$o \in O \tag{5.6}$$

$$N \subseteq M \tag{5.7}$$

In order to represent flow coefficient and flow balance in the constraint model, a two-dimensional set containing material unit pairs with coefficients is shown as:

$$\varepsilon \subseteq \{o \in O : (m, o) \in \varepsilon\} \tag{5.8}$$

The expression can be further explicitly split into two sets, representing outflows and inflows, with material nodes as points of reference. Hence, for each $m \in M$:

$$\varepsilon_m^{in} \subseteq \{o \in O : (o, m) \in \varepsilon\} \tag{5.9}$$

$$\varepsilon_m^{out} \subseteq \{o \in O : (m, o) \in \varepsilon\} \tag{5.10}$$

In addition, mutually exclusive groups $\{S_k\}_{k \in K}$ with $S_k \subseteq O$.

Model Parameters

In the model, the parameters are typically bounds and price for material nodes. Base bounds for non-product materials N will also include the value 0, to exhibit the disjunctive nature for the node. Similarly, capacity bounds, coefficient flow between unit pairs (vertices), fixed costs and proportional costs are parameters defined within the P-graph network nodes. Table 5.5 shows the specifications embedded in each P-graph component.

Table 5.5: List of parameters and bound constraints for each component.

Component	Specifications
Material Node	Type
	Price
	Lower Bound
	Upper Bound
Operating Unit Node	Lower Capacity Multiplier
	Upper Capacity Multiplier
	Proportional Cost
	Fixed Cost
Flow	Source Node
	Destination Node
	Flow Coefficient
Mutual Exclusion	Operating Unit Nodes involved

Decision Variables

The decision variables represent quantities which are solved by the model once it is optimised. The values can be continuous or binary in this case. Table 5.6 outlines the variables defined in each component.

Table 5.6: Model components, specifications, and associated variable types.

Component	Specification	Data Type
Material	Flow Value	Continuous
	Active	Binary
Operating Unit	Throughput	Continuous
	Active	Binary
Flow	Flow Value	Continuous
	Active	Binary

Model Constraints

The constraints consist of two categories: linear constraints for the MILP and GDP, both of which are outlined in this section. The model constraints with supported nodes can be visualised in Figure 5.14. In the linear constraints section, two main equation functions, which are Equation

(5.11) and Equation (5.12) are included in the Pyomo framework to emulate P-graph combinatorial network synthesis behaviour. In essence, the flow-to-capacity relation is shown as:

$$f_{mo} = b_{mo} q_o \qquad \forall (m,o) \in \varepsilon \tag{5.11}$$

$$f_{om} = b_{om} q_o \qquad \forall (o,m) \in \varepsilon \tag{5.12}$$

These equation functions represent the weightage assigned for material nodes entering the operating unit. The second equation calculates the flow when it exits the operating node. Note that the weightage or flow coefficient b_{mo} and b_{om} multiplied based on the capacity multiplier of operating units for both equations.

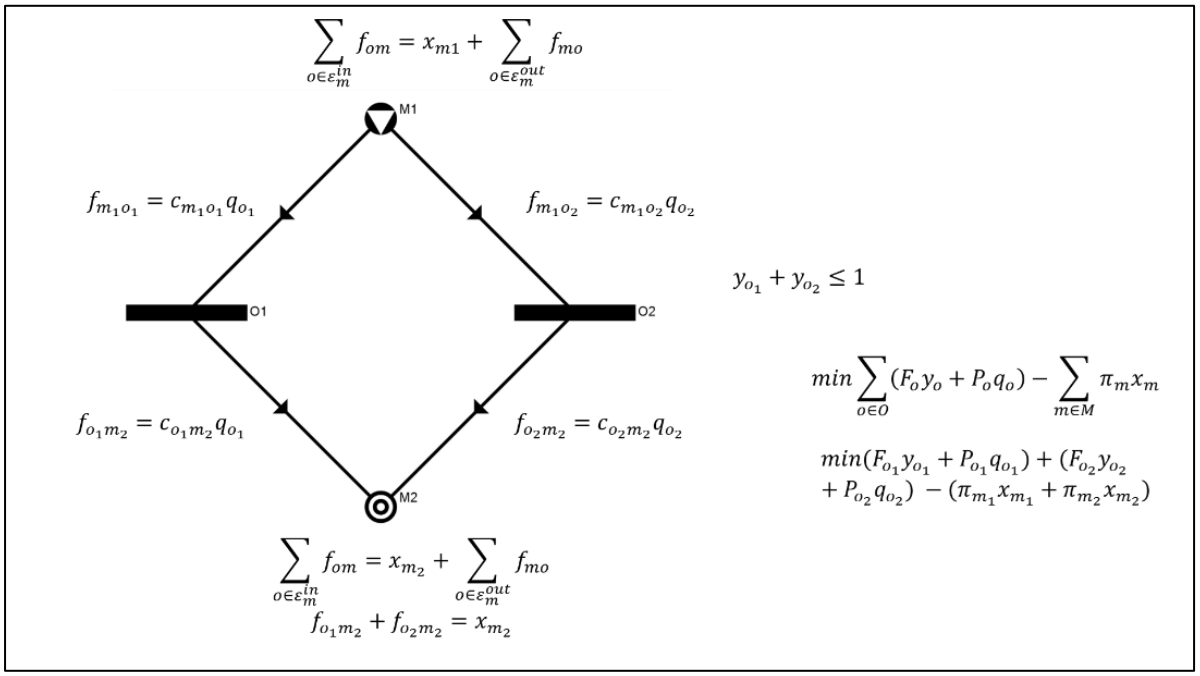


Figure 5.14: P-graph structure represented using an equation-based approach.

The flow coefficient equations are defined, setting the precedence for material flow balance along the process network model. The material flow balance calculates the values of the material node, which are accumulated or remained within itself, are joined by a series of incoming edges and outgoing edges which contain flow. Equation (5.13) for material flow balance is presented as follows:

$$\sum_{o\in\varepsilon_m^{in}} f_{om} = x_m + \sum_{o\in\varepsilon_m^{out}} f_{mo} \qquad \forall m \in M \tag{5.13}$$

where x_m is the net material term or accumulation in material. Based on the previous section, GDP disjunctions are used in the energy model which serve as ON and OFF indicators for the defined components, making sure up to or at least one indicator is active, depending on the situation. In the detailed model representation, the disjunctive behaviour will be attributed to three components in the model.

Generalised Disjunctive Methods

For the non-product material, each $m \in N$ introduces two disjuncts, with the Boolean variable of TRUE, indicating the disjunct is active or FALSE, indicating the alternate disjunct is active. The indicators $Y_{m0}, Y_{m1} \in \{0,1\}$ represent whether which constraints to follow in the model if one disjunct is enforced.

$$\bigvee_{i\in\{0,1\}} \begin{bmatrix} Y_{mi} \\ i = 0: x_m = 0 \\ i = 1: x_m^L \le x_m \le x_m^U \end{bmatrix}, \quad Y_{m0} + Y_{m1} = 1 \quad \forall m \in N. \tag{5.14}$$

The disjunction in operating units applies the same principle, with indicators $Y_{o1} + Y_{o1} \in \{0,1\}$ enforcing the disjunctions between a range of values and zero. With the expression $Y_{m0} + Y_{m1} = 1$, it sets the rule where one or the other should be applied to make the model feasible.

$$\bigvee_{i\in\{0,1\}} \begin{bmatrix} Y_{oi} \\ i = 0: q_o = 0 \\ i = 1: q_o^L \le q_o \le q_o^U \end{bmatrix}, \quad Y_{o0} + Y_{o1} = 1 \quad \forall o \in O. \tag{5.15}$$

Mutually exclusive operating units are an important feature embedded within the P-graph theoretical structure. This property must be explicitly defined, unlike mutual inclusion of operating units which can be represented by joining them within the network. The feature of mutual exclusivity cannot be imposed through such structural linking. Therefore, the feature requires explicit logical definition to ensure that only one of the units can exist in the P-graph solution structure. In the Pyomo equation, this P-graph property can be viewed as at most one operating unit existing in the solution for each given group S_k. Hence, the indicator for the disjunct is denoted as $y_o \in \{0,1\}$ which is equivalent to Y_{o1}.

$$\sum_{o\in S_k} y_o \le 1 \quad \forall k \in K \tag{5.16}$$

where $y_o = 1$ means the ON disjunct for unit o is selected, so q_o must satisfy $q_o^L \le q_o \le q_o^U$. $y_o = 0$ means the unit is OFF, so $q_o = 0$.

Based on the binary indicator Y_{o1}, it is shown for clarification for Y_{o0} as:

$$Y_{o0} + Y_{o1} = 1, \quad Y_{o0}, Y_{o1} \in \{0,1\} \tag{5.17}$$

Objective Function

The objective function is designed to return the global minimum costs for the solution structure, thereby minimising the total cost required to support decarbonisation of process heat configuration on a site-level resolution. The objective function is represented by a series of cost functions corresponding to the material and operating unit nodes. The first term represents the cost function by the operating units and material nodes for the second term. This is denoted in Equation (5.18):

$$min \sum_{o \in O} (F_o y_o + P_o q_o) - \sum_{m \in M} \pi_m x_m \quad (5.18)$$

where F_o is the fixed cost of operating unit, P_o is the proportional cost of operating unit, π_m is the price of the material unit nodes. Operating unit nodes which incur costs such as marginal levelised cost for existing generations, new electricity generation builds, capacity for heating technologies, electricity delivery costs, retrofitting coal boilers and methanol plant manufacturing cost. For the material nodes component, it consists of the cost incurred for resource usage, in this instance, such as biomass use. Material revenue is also obtained from deriving the second term of the equation, presenting profits earned from supplying end-use energy demand and green methanol sales.

Available Solvers

A multitude of optimisation solvers are available in the Pyomo environment. Pyomo provides dedicated interfaces for some solvers such as Branch-And-Reduce Optimization Navigator (BARON), COIN-OR Branch and Cut Solver (CBC), CPLEX, and Gurobi so as to make solvers compatible with the environment. Generic functions from Pyomo are also developed to call any solver that can read AMPL and write files into GAMs format and return the results back to Pyomo.

Table 5.7 lists the solvers that are installed, available, and tested within the Pyomo framework. However, one of the caveats when solving the P-graph problem while maintaining the structural rules and constraints defined by the P-graph framework, only MINLP and MILP solvers are employed to solve the model. Therefore, certain solvers, namely the COIN-OR Linear Programming Solver (CLP) and the Interior Point Optimizer (IPOPT), are included in the test as available solvers but do not have the capability to handle discrete decision variables. In addition, Bonmin solver is excluded from the tests as it is a local MINLP solver, which serves no purpose in a global MILP problem.

Table 5.7: Optimisation solvers tested by author in the Pyomo environment, including supported problem classes, licensing, and implementation details.

Solver	Mathematical Optimisation	License Type	Source	Remarks
GLPK	LP, MIP	Open source	(Makhorin, 2012)	ANSI C
Gurobi	LP, MILP, NLP, MINLP	Commercial	(Gurobi Optimization, LLC, 2026)	C
HiGHS	LP, MIP, NLP	Open source	(Huangfu and Hall, 2018)	C++
CLP	LP	Open source	(Lougee-Heimer, 2003)	C++
IPOPT	NLP	Open source	(Lougee-Heimer, 2003)	C++
CBC	MILP	Open source	(Lougee-Heimer, 2003)	C++
BONMIN	MINLP	Open source	(Lougee-Heimer, 2003)	C++
COUENNE	MINLP	Open source	(Lougee-Heimer, 2003)	C++

5.6.1 Cost Calculation for Components

Following the definition of the Pyomo model structure, the economic representation of technologies is introduced to determine the total annualised cost of components to be applied in each modular component. Each modular component in the framework is associated with cost parameters that represent capital investment and operational expenses. These costs include equipment purchase costs as well as additional project costs such as piping, controls, installation, and other direct and indirect expenses.

In this section, the cost equations used to quantify technology investment and operational costs are presented. These equations form the basis for the economic comparison of alternative technologies within the optimisation model and are incorporated into the objective function to determine the minimum-cost system configuration.

Capital recovery factor

The capital recovery factor is a ratio used to calculate the ratio of the constant annuity costs to the total present value of the costs for a set amount of time (Jenkin *et al.*, 2019).

$$CRF = \frac{WACC(1+WACC)^n}{(1+WACC)^n - 1} \quad (5.19)$$

where $WACC$ is the Weighted Average Cost of Capital (WACC) for different components would be different as it varies based on industries. For example, the WACC for grid assets is 6%, 3% for heating technologies and other components would be 4% for industry and 7% for new technologies.

Annualised capital cost

The equation to calculate annualised capital cost, which is the cost that occurs equally annually over the project lifetime (HOMER Energy, 2016). The annualised cost is given in Equation (5.20):

$$A_{cap} = C_{cap} \cdot CRF \quad (5.20)$$

Annualised fixed operating cost

The operating cost is expenses associated with the fixed operating expenses and maintenance costs of the components (HOMER Energy, 2016). It is presented in Equation (5.21) as:

$$A_{op,fix} = C_{cap} \cdot f_{om} \quad (5.21)$$

where f_{om} is the ratio of fixed operating and maintenance costs.

Variable operating cost

The variable operating cost is comprised of prices incurred from raw materials, fuel and other commodities. The variable is given by $A_{op,var}$.

Total annualised cost

The total annualised cost represents the annualised capital costs and operating costs.

$$TAC = f_{ins} \cdot A_{cap} + A_{op,fix} + A_{op,var} \quad (5.22)$$

Once solved, the resulting outputs must be standardised into a format suitable for interpretation and post-processing.

5.7 Results Standardisation and Post-Processing

The results from the optimisation run are presented in Figure 5.7. Each worksheet displays its corresponding outputs based on the boundaries and parameters associated with a specific identifier. The identifier serves as the key reference for determining the layer, component, and time step, where applicable, to which each variable belongs within the energy model. Outputs generated from both the P-graph and Pyomo approaches are standardised into the same format in the results section. This consistency supports seamless interchange between P-graph-based and Pyomo-based optimisation, which is particularly useful for sequential optimisation and post-optimisation stress testing, where key parameters are adjusted to evaluate the response of the optimised system structure.

The primary motivation for adopting this output format is to facilitate efficient extraction and aggregation of results. By organising outputs according to identifiers and attributes, results can be readily analysed by system layer, technology type, spatial grouping, or other user-defined criteria, enabling clear and meaningful interpretation of model outcomes. Several analysis modules are included in the framework, while additional customised analysis can be carried out by the modeller. The general concept of operating the post-processing tools is as displayed in Figure 5.15.

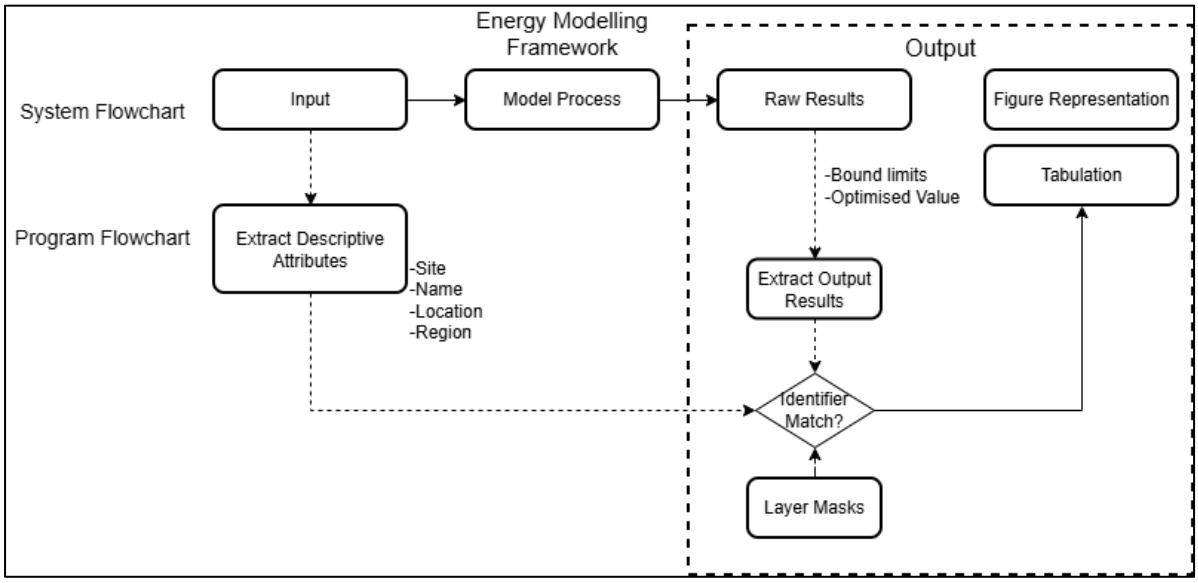


Figure 5.15: Conceptual design and rationale for presenting post-processed outputs with descriptive attributes.

The input database is first integrated into the process model to enable simulation of the energy system. In addition to the attributes required directly by the model, supplementary attributes that are not used in the optimisation are retained to support the identification of additional features and patterns within each system layer. Once the model runs successfully, the output results are automatically extracted and saved to a designated directory within the energy modelling environment.

The representation of different layers, components, and network connections applied in the energy model is analysed through the application of filtering techniques based on the Boolean method and logical masks. These masks implement string pattern-based filtering conditions on identifier codes to correctly classify nodes according to their respective categories. Logical filtering is applied to the identifier strings to select nodes belonging to specific layers or component types. When identifiers are matched consistently, model parameters and output results can be combined in the same row, allowing results to be aggregated and processed more easily. Table 5.8 represents a template embedded in the energy model, where identifier strings are matched by length, prefix, and unique masks to identify key P-graph nodes within the Pyomo optimisation framework. Once the results data has merged with the input spreadsheet, figures and tables can be done by categories. This can be done by pivoting categories based on their capacity, location, POC connection points, region, type, and many more. This analysis can be performed by the user because the output results are exported in spreadsheet format.

Table 5.8: Naming convention used to extract attributes from reference layer identifiers from the dataset.

Reference Layer	Specification	Length	Prefix	Slice Position (Python str.slice)
Factory (Heating Tech)	Flow	10	O3	5, 8
Generation	Capacity	6	O14	:3
Methanol Plant	Capacity	11	O10	6:9
Biomass Connection	Flow	11	O8	7
Biomass (Type)	Resource	7	M8	6
Sub-transmission	Capacity	9	O18	3:6
Transformer	Capacity	8	O	-2:
Transmission Grid	Capacity	12	O6	-3:-1
Distribution Lines	Flow	10	O3	5:8

5.7.1 Post-processing for the Factory Example

Each identifier contains a serial number that uniquely represents a node within the P-graph framework. A generic factory input database includes attributes such as index, industry type, process type, company name, coordinates, connected POC, region, energy carrier, current primary fuel, and other relevant characteristics. These attributes provide contextual information

for post-processing and analysis, even if their technological pathways are not selected in the optimisation. Other reference layers are available in Appendix B.

Filtering requirements are based on structured identifier patterns. The prefix indicates the node type and system layer, where the character O denotes an operating unit node and the value 3 identifies the factory layer. Identifiers with a length of ten characters represent operating units that are time dependent. Within the factory layer, the unit flow number is encoded in positions five to eight of the identifier string. This stream number segment distinguishes between different operating unit nodes, including those representing heating technology flows in the factory site.

The mid-section of the identifier string is used to represent the type of heating technology associated with each flow. These technology codes are stored in a JSON format to allow clear mapping between identifier strings and technology names and to facilitate future modification. The technology codes used in this study are listed below:

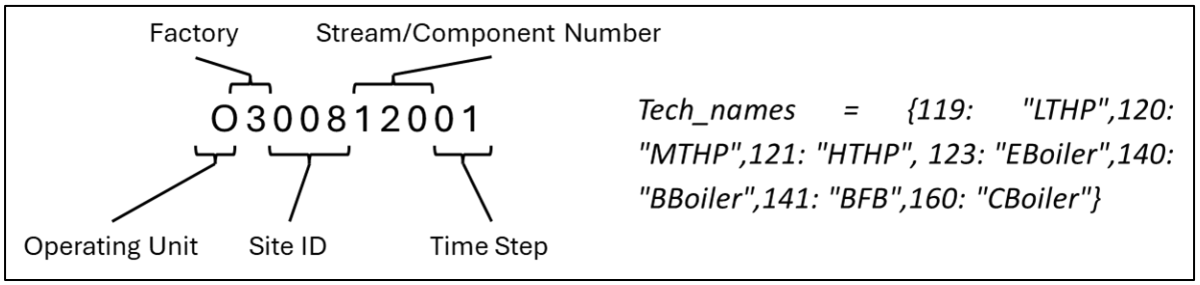


Figure 5.16: Identifier for the energy input to the respective heating technologies.

Future extensions of the model may include additional datasets such as proposed installed capacity at the site level. In this case, the identifier structure would be modified. Because installed capacity is not time dependent, the identifier length is reduced to eight characters. The technology codes associated with capacity representation are shown below:

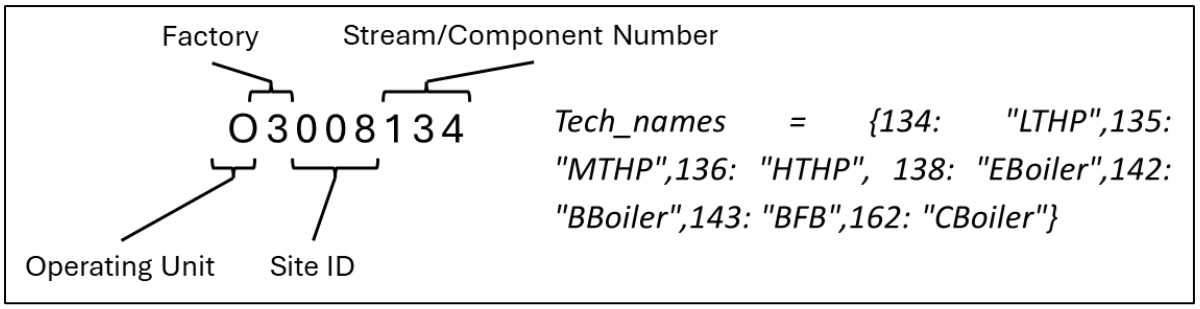


Figure 5.17: Identifier for the capacity sizing for the respective heating technologies.

5.8 Illustrative Example of the Framework and Modular Operations

An illustrative example is presented in Figure 5.19 using a simplified system consisting of a single electricity generator operating with a single biomass resource at one location, together with a single factory whose heat demands span multiple temperature levels. The model is configured

using a monthly time resolution with a single time step. As the final part of Section A, this example serves to demonstrate how the overall framework and modular operations function in practice before the case study is presented.

In this example, a proposed run-of-river hydroelectric power station is included. Its utilisation factor is assumed to be one, resulting in a monthly energy multiplier of 730 to reflect the number of operating hours in a month. The maximum allowable installed capacity of the hydropower station is set at 1000 MW. The contingency ratio is enforced in the example as part of the generation layer, as this ensures no power shortages due to peaking. The ratio is set at 0.6 in the example, representing the average demand over installed capacity.

The example contains two POCs and one factory, each with its respective electricity demand. The demand for POC A is 40 GWh per month, whereas POC B is 10 GWh per month. Examination of the resulting solution structure shows that the model adheres to the mutual exclusion, or disjunctive, rule between paired operating units. All nodes responsible for transferring energy across the value chain operate and satisfy the imposed boundary conditions. As a consequence, electricity flows only in the direction as indicated by the arrows.

At the distribution level, the POC nodes are branched into two categories: nonindustrial electricity demand and industrial process heat demand within the factory. Electricity costs in the model are proportional to the electricity price assigned to each POC or the factory negotiated price, which is set at 142 NZD MWh^{-1} for this example. Therefore, the cost of electricity is determined based on the quantity of electricity consumed as an input fuel rather than on the end-use thermal output, due to the wide variation in efficiencies among electrified heating technologies.

Within the factory module, the electricity supply node splits into multiple pathways associated with different heating technologies that serve the five temperature levels included in the example. A parallel pathway arises from the biomass resources, namely non-woody biomass and woody biomass priced at 20 NZD GJ^{-1} and 10 NZD GJ^{-1}, which both require their respective biomass heating technologies for energy generation. As the thermal output quality from these biomass systems is not dependent on the temperature level, the energy nodes diverge only after the heat has been generated. The biomass heating technologies represented in the model include a biomass boiler, a bubbling fluidised bed reactor, and a retrofitted coal boiler. However, only woody biomass is compatible as a fuel input for the retrofitted coal boiler due to the physical properties of the solid fuel. The biomass resource is assumed to be located 100 km away from the factory.

The factory demand for decarbonisable process heat is set at 1 GWh per month for all temperature levels. Each product node receives heat through pathways originating either from electrification technologies or from biomass-based technologies, depending on the configuration selected by the model.

When applying the Solution Structure Generation (SSG) algorithm from the P-graph framework to the example consisting of a single factory, two biomass types from the same location, and one electricity generation site, the model identifies 27,120 possible solution structures. The SSG

algorithm enumerates all combinatorially feasible structures that can be derived from the maximal structure without considering the node level boundary constraints. Since the problem size in this example is fairly small, it is possible to examine all feasible optimal solution structures explicitly. After applying the accelerated branch and bound algorithm, the model identifies 94 feasible solutions, obtained after evaluating 1,011 subproblems within the P-graph network. Figure 5.18 presents the optimal cost associated with each feasible solution.

By examining the permutations of heating technologies configurations in the model, there are a total of 94 feasible solution structures for each industrial site, assuming each site has demand in all temperature bands. Without accounting for the number of generation sites, the number of biomass connections, and transmission links, it is assumed that all 298 major industrial sites operate independently and that each site has heat demand across all temperature levels. Under this assumption, the total number of possible factory level modular system combinations is 94^{298}. This reflects the full combinatorial space in which every site can independently adopt any one of its 94 feasible configurations. The number of possible configurations will inherently increase when individual biomass resource to factory connections are established throughout the energy model.

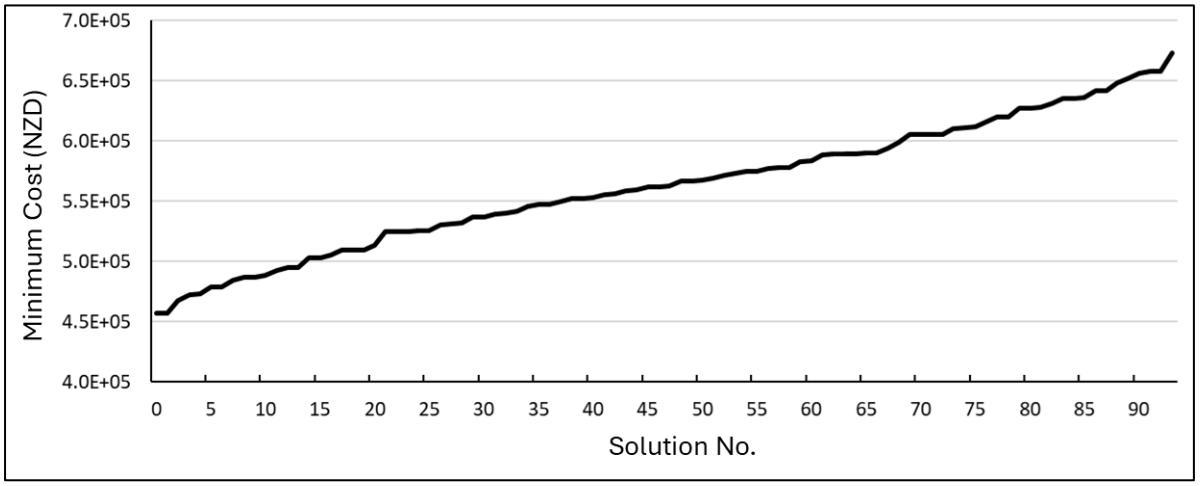


Figure 5.18: Ranking of all feasible solutions by total cost.

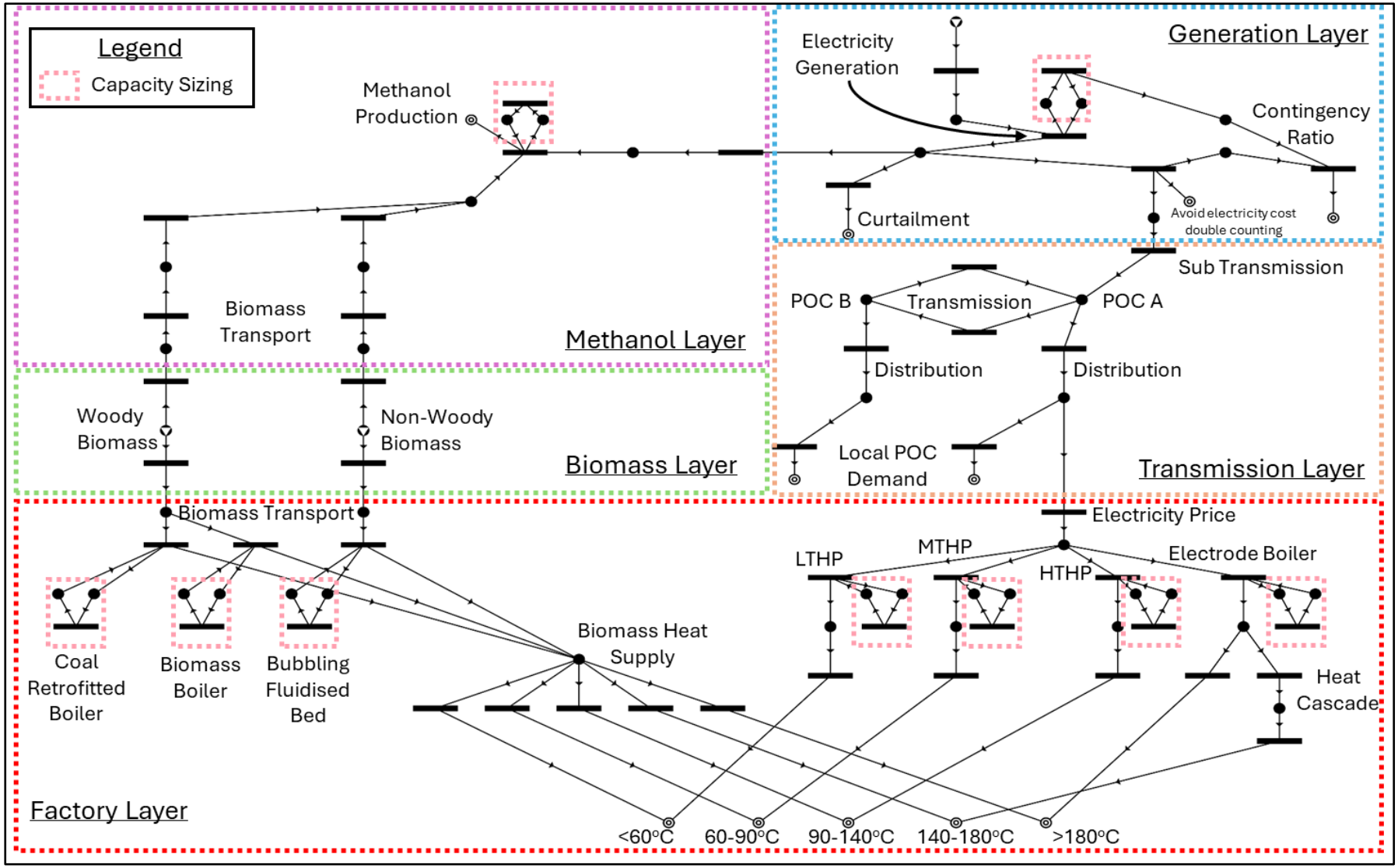


Figure 5.19: Single time-step system with one generation, factory, biomass, and methanol unit.

The total cost of the system is 456,636 NZD, with 243,836.6 NZD from biomass and 212,799.63 NZD from the electricity side. The capacity of hydroelectricity generation is set at 74.3 MW. However, the minimum hydroelectricity generation required is 44.55 MW to supply the industrial and local demands, while the remainder is reserved for contingency. In total, 54.2 GWh per month of electricity is generated to meet the demand.

Looking at the best solution structure, LTHP, MTHP, HTHP and CB are implemented as the selected technologies. At the distribution end, 1.17 GWh of electricity is delivered to the factory site, where electricity is supplied separately to different heat pumps for different temperature levels. However, for higher temperature levels such as 140 to 180°C and above 180°C, heat is supplied by CB. In this case, a total of 2.71 GWh per month of biomass is supplied to the factory site, contributing 2.03 GWh per month to process heat after accounting for losses. The biomass supplied to the factory is woody biomass. As a summary, the heating technology capacities are shown in Table 5.9 below.

Table 5.9: Installed heating technology capacities for the best optimal solution structure.

Technology	**LTHP**	**MTHP**	**HTHP**	**CB**
Capacity (MW)	0.37	0.64	0.99	4.7

In a small-sized energy problem, it is possible to evaluate the most cost-effective configurations involving electricity only, biomass only, and hybrid electricity and biomass systems. Table 5.10 summarises the technology configurations and associated costs for these cases. The cost distribution across these solutions indicates a strong preference for electrification of industrial heating. All 10 best structures employ electrified heating technologies, with heat pumps playing a central role.

In contrast, the 94th best solution corresponds to a fully biomass-based heating system, in which no heat pump driven heating technologies are used. In this example, the configuration performs significantly worse from a cost perspective, which demonstrates that a factory node relying entirely on biomass is not economically favourable under the specified conditions.

When methanol production is set at 100 tonnes per month, the model shows an increase in electricity generation capacity from 74.25 MW at zero methanol output to 74.85 MW. The cost of electricity produced is represented by the flow value present at the product node. Methanol production is assumed to require 436 MWh of electricity.

On the biomass side, 396 MWh of forest residues are transported over 100 km to the generation site. The transport cost is 396 NZD for the 100 km distance, and a resource and harvesting cost of 14,256 NZD is included in the model. The fixed operating and capital cost for the methanol plant is set at 700 NZD t^{-1}. Under these assumptions, the levelised cost of green methanol production in this example only shows a profit when the methanol price reaches 1,000 NZD t^{-1}.

Table 5.10: Summary of selected feasible heating technology configurations and associated system costs.

No. Sol	Heating Technologies Used	Electricity Input (GWh)	Biomass Input (GWh)	Electricity Cost	Biomass Cost	Factory's System Cost
1	LTHP, MTHP, HTHP, CB	1.17	4.65	212,800	243,837	456,636
2	LTHP, MTHP, HTHP, BFB	1.17	4.65	212,800	244,148	456,948
3	LTHP, MTHP, HTHP, BB	1.17	4.36	212,800	254,354	467,153
4	LTHP, MTHP, CB	0.59	6.95	107,685	364,830	472,515
5	LTHP, MTHP, BFB	0.59	6.95	107,685	365,296	472,980
6	LTHP, MTHP, BB	0.59	6.52	107,685	380,566	488,250
7	LTHP, MTHP, HTHP, EB, CB	2.18	2.32	357,230	121,609	478,838
8	LTHP, MTHP, HTHP, EB, BFB	2.18	2.32	357,230	121,764	478,994
9	LTHP, MTHP, HTHP, EB, BB	2.18	2.17	357,230	126,854	484,084
10	LTHP, HTHP, CB	0.80	6.94	144,958	364,223	509,181
94	EB, BB	2.07	6.45	296,462	376,741	717,022

**LTHP = Low-Temperature Heat Pump, MTHP = Medium-Temperature Heat Pump, HTHP = High-Temperature Heat Pump, EB = Electrode Boiler, BFB = Bubbling Fluidised Bed Reactor, CB = Retrofitted Coal Boiler.*

5.8.1 Changing the Time Step

Increasing the number of time steps in the model does not inherently increase the number of feasible solution structures generated by the P-graph framework. However, the number of constraints applied increases the number of variables governing the system boundaries of the model (Figure 5.20). The number of structural configurations therefore remains largely unchanged compared with the single time-step case study.

Introducing multiple time steps expands the size of the data tables and increases the number of operational variables, thereby increasing the solving time. However, each time step is constrained to its own set of operating units, meaning that temporal variation is captured within the operational layer rather than by creating additional structural pathways. As a result, the time steps primarily represent changes in energy flows over time rather than modifying the underlying combinatorial structure of the network. Consequently, the computational burden associated with structural enumeration does not scale with the number of time steps. The performance summary for the case study with different time steps is presented in Table 5.11.

Table 5.11: Model size and computational characteristics for different numbers of time steps.

Time Step(s)	1	2	6	9	12
Size Number of Constraints	697	1,243	3,427	5,065	6,703
Number of Variables	446	801	2,221	3,286	4,351
Number of Binary Variables	202	354	962	1,418	1,874
Number of Integer Variables	202	354	962	1,418	1,874
Number of Continuous Variables	244	447	1259	1,868	2,477
Total Time	4.08	4.93	6.98	7.97	9.31

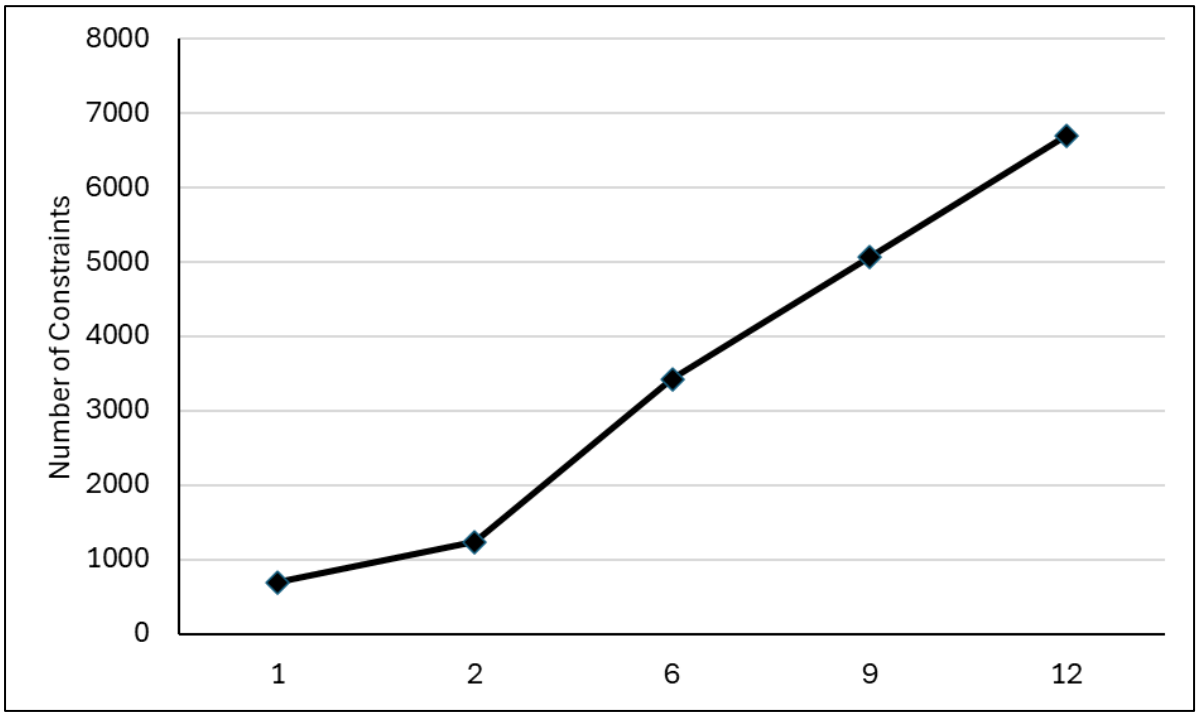


Figure 5.20: Modelling summary showing the relationship between the number of constraints and the number of time steps.

5.9 Chapter Summary

This chapter has translated the structural architecture developed in Chapter 4 into a mathematical and computational modelling framework, forming the operational backbone of the energy system model. It bridges the gap between conceptual superstructure design and executable optimisation by formalising the relationships between components, variables, and constraints.

First, the chapter outlined the data handling and input structure, establishing how model parameters, component attributes, and system configurations are imported and standardised. This ensures consistency in representing diverse datasets across multiple system layers, including generation, industrial demand, biomass supply, and network infrastructure.

A key contribution of this chapter is the development of the integrated P-graph and Pyomo workflow, where the superstructure generated through P-graph is translated into an algebraic optimisation model. The introduction of a P-graph-to-Pyomo porting tool enables automated conversion of structural relationships using GDP formulation. This approach preserves the transparency of P-graph while leveraging the high computational strength and flexibility of Pyomo for solving large-scale optimisation problems. The use of GDP allows explicit modelling of discrete technology choices and logical relationships, which are essential for capturing fuel-switching decisions and infrastructure investments. These formulations are subsequently reformulated into solvable MILP problems, ensuring computational tractability while maintaining model fidelity.

In addition, the chapter also introduced result standardisation and post-processing procedures, enabling outputs to be structured in a consistent format for interpretation, comparison, and visualisation. This is done by masking the identifier of the node representing each layer in the superstructure.

A short illustrative case study was presented to demonstrate the functionality and validity of the modelling workflow, highlighting how the framework can generate feasible system configurations and serving as an example before progressing to the large-scale problem. Building on this implementation, Chapter 6 applies the framework to a real-world case study of New Zealand, where detailed system inputs, spatial data, and sector-specific characteristics are introduced.

Section B: Case Study Workflow and Comparative Analysis of Simulation Modes

In Section B, the focus shifts from the development of the energy systems framework to the case study and simulation modes used to apply the modelling framework. Chapter 6 introduces the New Zealand case study. Extensive data reconciliation and preprocessing were undertaken to standardise datasets from multiple sources, and the key assumptions adopted in the case study are also presented.

One of the notable contributions of the case study is the integration of spatial tools with the P-graph-Pyomo energy modelling framework. The reason behind the implementation of GIS tools is to add spatial information to original raw datasets to be designed for spatial optimisation. Spatial information is incorporated not only for the studied sites, but also for the connections between locations across different layers to account for transport and infrastructure costs. In addition, location-dependent intermittent electricity generation namely wind and solar are passed to the Renewables.ninja tool to extract the generation data. At the end of Chapter 6, the model inputs for the New Zealand case study undergoes a series of consistency checks with measured data. After that, the next step is to demonstrate how different optimisation formulations can be used, depending on the study objective and the limitations of available data.

Three simulation modes are presented across Chapters 7, 8, and 9. Chapter 7 introduces a single deterministic multiperiod optimisation, which provides a baseline least-cost fuel-switching outcome under a fixed set of assumptions. Chapter 8 extends this approach using sequential deterministic optimisation to reflect staged decision-making and to incorporate additional system structure that is difficult to represent in a single solve. Chapter 9 then introduces stochastic optimisation to test the robustness of decarbonisation pathways under variability and uncertainty.

Together, these simulation modes show how the same generic energy modelling framework can be applied flexibly for different purposes, ranging from high-level pathway screening to more operationally realistic assessments, while making transparent what is captured in each formulation and what is deferred when data or model scope limits apply.

Chapter 6: System Definition and Context for New Zealand

6.1 Introduction

In this chapter, a New Zealand case study is introduced to demonstrate the workflow used to assess energy system design under high spatial resolution. The chapter outlines the national context needed to define the key parameters, decision variables, and constraints represented in the model. It then presents the full workflow up to model generation, including preprocessing steps such as geospatial processing, distance calculations, and electricity generation profile construction. The resulting model inputs integrate industrial process heat demand, electricity generation, biomass resources, and associated cost assumptions. At the end of the chapter, the pre-processed input data from the preprocessing tools and the underlying assumptions are compared to measured values to make sure the model inputs are correct. This chapter establishes the case-study inputs used to demonstrate the framework's capability. The modelling choices here are New Zealand specific and not part of the generalisable methodological contribution.

Data extraction and preprocessing compile information from multiple sources, including government databases and selected private datasets. Transformation steps are applied to standardise formats and synchronise inputs before integration. Where required, synthetic datasets such as meteorological reanalysis-based generation profiles are generated. Spatial datasets are enriched by assigning coordinates and estimating distances or connectivity between relevant locations. These steps convert raw datasets into model-ready parameters that represent both system components and their relationships.

6.2 Scope and Modelling Dimensions

The New Zealand case study is structured around the five key modelling dimensions which were identified from the energy systems literature in Chapter 2. These are space, time, component, energy transformation and scenario. As the energy modelling framework is designed to solve high-fidelity energy system problems, the spatial resolution is represented by geocoded industrial sites, generation sites, biomass locations, and their assigned grid connection nodes, with cost distance-based relationships used to approximate transport and connection costs.

Time is represented through processed time series inputs for electricity demand and renewable generation, with temporal aggregation applied to produce multiperiod time resolution while retaining seasonal and intermittent patterns. Component refers to the modular representation of system elements, where factories, grid nodes, generation assets, biomass supply points, and conversion technologies are treated as discrete objects that can be combined consistently within the broader model structure.

Energy transformation is represented through technology options that convert primary energy carriers into useful heat or electricity services, with technology feasibility or even sustainable

green methanol products sourced from biomass steam methane reforming and electrolysed green hydrogen.

Scenarios are represented through alternative parameter settings and boundary assumptions, allowing comparison of system configurations under different constraints or future conditions. As electricity price, cost of biomass and availability of biomass are dependent on the time dimension, they are discussed in Section B of the thesis.

Within this case study, the process heat transition problem is scoped to indirect heat applications where conventional technology substitution is feasible, particularly systems using steam or comparable heat transfer media. High-temperature, highly process-specific applications such as kilns, reformers, and electric arc furnaces are excluded.

The reference year for system-wide input compilation and validation is set to 2022, which is used to align national totals and benchmark electricity demand and generation profiles. However, site-level industrial process heat inputs are taken from the most recent available year, 2024, because there is no consistent site-level dataset are available for earlier years. The future projected year is set as 2037, which is aligned with the year when government policy to phase out all existing industrial coal boilers (Ministry for the Environment, 2024).

6.3 Data Sources and Reconciliation Issues

The New Zealand case study compiles inputs from multiple sources to represent industrial process heat demand, electricity generation and grid infrastructure, biomass availability, and the spatial relationships that link the entries between layers or within the same layer. Most original raw datasets were produced for reporting or administrative purposes rather than spatial optimisation, so the formats differ in scope, definitions, identifiers, and completeness. Therefore, data reconciliation and preprocessing are required before the information can be converted into the format where it is compatible to be implemented as parameters.

On the industrial process heat side, the Regional Energy Transition Accelerator (RETA) Database (EECA, 2025b) database and Heat Plant Database (Bioenergy Association of New Zealand, 2011) characterise industrial heat using different approaches. The RETA Database focuses on onsite allowable process heat that can be decarbonised based on fuel-switching potential, whereas the Heat Plant Database characterises process heat based on boiler type and installed capacity. These differences lead to inconsistencies in site lists, sector totals, and the interpretation of what constitutes eligible heat demand. To address this, the compiled dataset is first assessed at a high level by comparing aggregated sector totals with other agency datasets, then progressively narrowed by applying consistent inclusion rules, including filtering by heating method and temperature level and removing sectors outside the study scope.

On the electricity system side, multiple sources are required because no single database provides a complete and consistent representation of generation assets, grid nodes, and network constraints. Electricity generation data are compiled from the Electricity Authority and supplemented by other sources such as Open Infrastructure Maps and local power companies

because some datasets aggregate schemes (particularly hydro) rather than listing plant-level units. Such data process requires additional compilation and cross-referencing so that generation sites, capacities, year built, commissioning status, and operator information are represented consistently.

Grid-related datasets also show inconsistencies. Information on POCs from the Electricity Authority does not fully align with Transpower reports. This mismatch reflects redundant or phased-out points remaining in some lists, as well as planned or proposed nodes appearing in Transpower material. In addition, transmission security information is not fully consistent across sources. To represent grid adequacy and constraints, n-security limits are compiled using numeric inputs from both ERGO reports (EECA, 2025c) and Transpower's database (Transpower New Zealand, 2023a), with reconciliation required when definitions or coverage differ between both sources.

Biomass resource inputs are compiled from multiple sources and require both spatial and non-spatial reconciliation. The spatial extent of biomass availability is derived from a GIS layer (Ministry for the Environment NZ, 2024), while biomass quantities are taken from the report "Residual biomass fuel projections for New Zealand; 2021" (Hall, 2022). RETA provides cost inputs, which include differences between the North and South Islands (EECA, 2025c). Because land-use layers are polygon-based data and not directly usable in the optimisation structure, preprocessing is required to filter non-relevant land-use classes, remove minor polygons, decluster multipart geometries, and convert biomass areas into point representations via centroid generation. Spatial joins are then used to attribute administrative regions. These steps are performed in ArcGIS for spatial verification, while Python-based ArcPy is used for processing. The workflow converts raw land-use geometries into model-ready biomass supply nodes that can be linked to demand and generation sites through distance-based cost parameters.

A further limitation across industrial, electricity, and biomass datasets is that many entries are not supplied with coordinates. While tables often provide site names, capacities, and descriptive location fields, they are not natively suitable for spatial modelling. To address this, geocoding is applied to industrial sites and electricity generation sites so that each entry, such as place name, can be assigned coordinates and linked to spatial layers. These spatial attributes then support distance calculations, assignment of industrial sites to biomass centres, and thermal-powered generation sites to biomass locations.

In addition to spatial reconciliation, the case study requires time series inputs for electricity prices and demand, which are later reformulated into the FPVV proxy structure. High temporal resolution datasets from the Electricity Authority, specifically demand and wholesale electricity prices at 30-minute intervals, are too large to extract manually at a national scale and often result in server timeout errors when queried directly from the web interface. Therefore, automation is implemented using ChromeDriver and Selenium to download the required time windows iteratively and consolidate outputs into structured datasets suitable for aggregation and preprocessing. This facilitates the creation of uniform temporal price and demand profiles that

align with the modelling time structure. The extracted data is also pre-processed to match the time resolution defined in the model dimension.

Overall, data reconciliation consists of harmonising identifiers and naming conventions, standardising units and formats, enriching records with missing geospatial attributes, and resolving mismatches by applying consistent inclusion rules. These steps ensure that all inputs can be aligned into a unified set of parameters that represent both system components and the relationships among components. After data preparation, the user can import the predeveloped modular structures into the model by incorporating them into the Python script. Parametric settings for the model can be modified, and these changes propagate automatically through the structural network. This ultimately forms the final model-ready input dataset for the New Zealand case study.

6.4 Process Heat

6.4.1 Brief Overview of Process Heat in New Zealand

In New Zealand, process heat can account for around 30% of the total energy supply. Due to resource availability and geographic constraints, the North Island mainly uses natural gas, whereas the South Island has a larger availability of lignin-based resources and therefore relies more heavily on coal boilers (EECA, 2024). The sector that consumes the most process heat is dairy product manufacturing, followed by chemicals, metals, and wood products.

New Zealand uses 115.7 PJ of energy in industrial process heat (EECA, 2024). Several databases can be used to estimate process heat in New Zealand, and more recent data include the 2025 RETA Database's dataset, which provides site-level load demand for factories for decarbonisation assessment (EECA, 2025b). An important boiler survey is the 2011 Heat Plant Database completed by Scion, which includes boiler capacity, year installed, and fuel type for medium to large-scale enterprises (Bioenergy Association of New Zealand, 2011). Despite the information sheet being outdated, it is not considered redundant or obsolete, as it remains useful for identifying the locations of boilers, their capacities, and the types of fuel used, as well as for estimating the amount of energy supplied and determining which sites require process heat for their operations. Moreover, some sites preferred to keep their boiler specifications confidential during the survey leading to an incomplete yet still valuable dataset for informing process heat decarbonisation.

In this section, the compiled data is first assessed at a high level by comparing aggregated sector totals with other agencies. The scope is then progressively narrowed by filtering by heating method and temperature level, followed by removing sectors that are outside the study scope. Finally, the section provides the justification for the cut-off used to determine how many sites are included in the model inputs based on heat demand.

6.4.2 Major Industrial Sites with Process Heat

The Heat Plant Database and electricity generation datasets represent two distinct components in the energy systems model. The Plant Heat Database includes information about the boiler type and heat demand, as well as temperature profiles.

To estimate end-use process heat demand at the individual site level, the compiled dataset is compared with several studies produced by different agencies and government bodies. Energy reports from the Ministry of Business, Innovation and Employment (MBIE) namely the "Energy in New Zealand 2024" report (MBIE, 2025a) and the EECA Energy End-Use Database (EEUD) (EECA, 2024) are aggregated by sector and compared with the compiled data.

In Table 6.1, the compiled values represent fossil fuel use for heating and cooling. MBIE reports total energy use, while EECA reports heating and cooling energy use from fossil fuels only. MBIE values represent total energy use and therefore often include non-heating end uses such as motors and stationary equipment. In contrast, EECA separates heating and cooling from motors and other end uses.

Table 6.1: Comparison of sectoral energy fuel consumption (TWh) between compiled results and national datasets from MBIE (2024b) and EECA (2024) for the year 2022.

Sector	Compiled Version	(MBIE, 2024b)	(EECA, 2024)
	TWh		
Chemical	4.0	4.2	**6.6**
Dairy	8.7	12.2	5.4
Food	1.5		1.0
Meat	1.1		0.9
Metal	0.7	0.7	0.7
Non-Metal	1.0	2.0	1.2
Other	1.1		1.2
Pulp	2.5	7.5	9.8
Wood Processing	5.8		
Total	**26.4**	**25.8**	**26.7**

In Table 6.1, it is also worth noting that MBIE groups the dairy, food, and meat industries under a single food category. In addition, electricity consumption in the metals sector is excluded from the table because it is dominated by the aluminium smelter, which draws electricity from the Manapouri Power Station and is already electrified, and therefore largely decarbonised. Within the EECA EEUD energy summary, oil and gas and petrochemicals are combined within the chemical sector, which increases the reported energy use for that category. There is also a discrepancy in the wood processing category. This discrepancy is because the compiled database includes minor wood processing facilities as well as pulp and paper sites across New Zealand, whereas MBIE separates wood processing from agriculture, where agriculture includes planted crops. This difference in sector definitions can lead to differences in aggregated values. EECA includes black liquor, an intermediate product used in pulp and paper, within the total for the

sector, which further increases the reported value. As a summary, the compiled version is in good agreement with both datasets, and the evaluation based on temperature levels can be assessed in the following section.

6.4.3 Temperature levels

According to the EECA EEUD database, the process heat requirement is as follows: 6 TWh is below 100°C, 12.9 TWh is between 100°C and 300°C, and 13.2 TWh is above 300°C (Figure 6.1). However, within the scope of this model, only temperature levels associated with indirect heating are considered, particularly where high-pressure steam or lower is used as the heat transfer medium. As a result, technologies such as reformers, kilns, and electric arc furnaces which exist in New Zealand are excluded from the technology switching analysis.

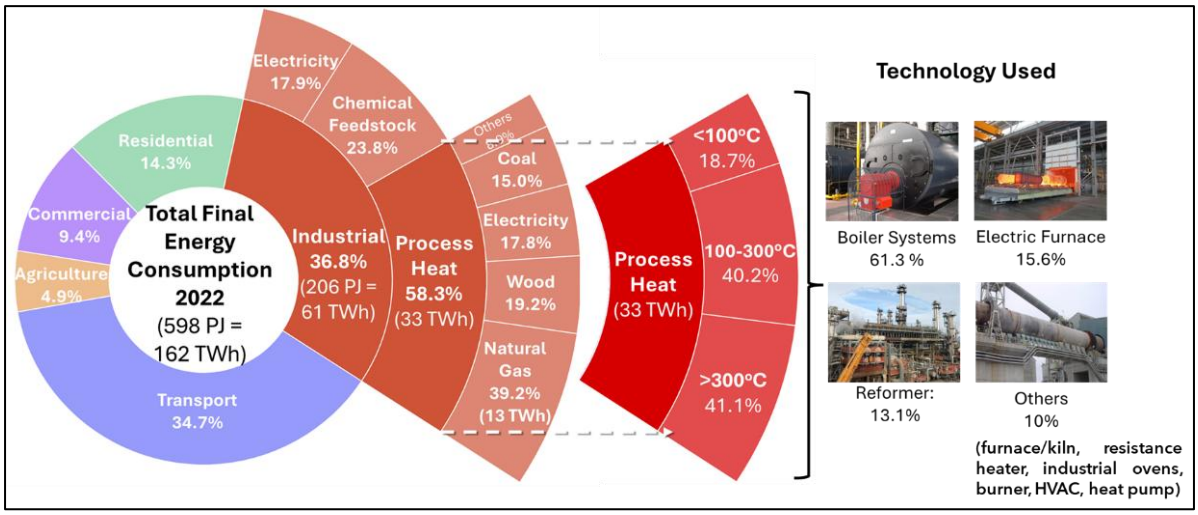


Figure 6.1: Industrial process heat demand in New Zealand, prepared by the author using data from EECA (2024).

More importantly, for these high-temperature applications, decarbonisation often requires more than fuel-switching, as it may involve changes to the underlying process or reaction pathway. In contrast, the focus of this study is on sectors where indirect heat provision and conventional technology substitution are more feasible. These include key contributors to the New Zealand economy such as dairy, meat processing, textiles, wood pulp and paper, non-metallic minerals, and the chemical sector.

Nevertheless, large energy users such as Glenbrook Steel, Tiwai Point Aluminium Smelter, Ballance Agri-Nutrients, and Golden Bay Cement are excluded from the analysis due to their high-temperature process requirements, which are beyond the scope of this study. Similarly, as geothermal energy is location dependent which is concentrated in the Bay of Plenty region, it is netted off from the site-level heat demand prior to optimisation. This is because the demand is already decarbonised and does not require a fuel-switching decision. In short, decarbonisation of process heat in New Zealand within this study is constrained to temperatures up to 500°C, reflecting the operational limits of the boilers and heat pump technologies considered. By removing furnaces, kilns and reformers from the total process heat supply, the amount that can be decarbonised is reduced down to 17 GWh per annum.

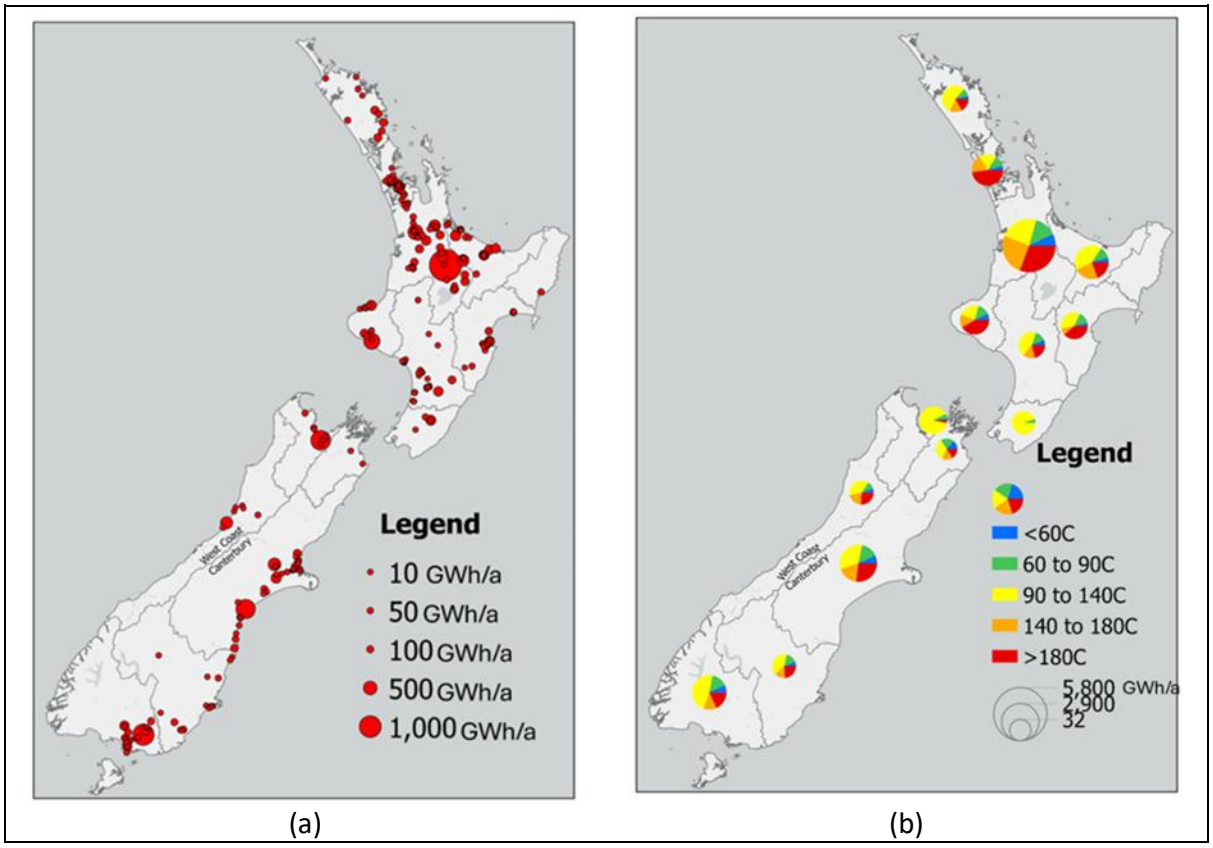


Figure 6.2: (a) Decarbonisable industrial process heat in New Zealand, compiled from RETA and the Heat Plant Database; (b) Spatial distribution of decarbonisable process heat temperature levels, based on Table 4.2 and European Commission (2008), supplemented by referenced and partly confidential industrial data.

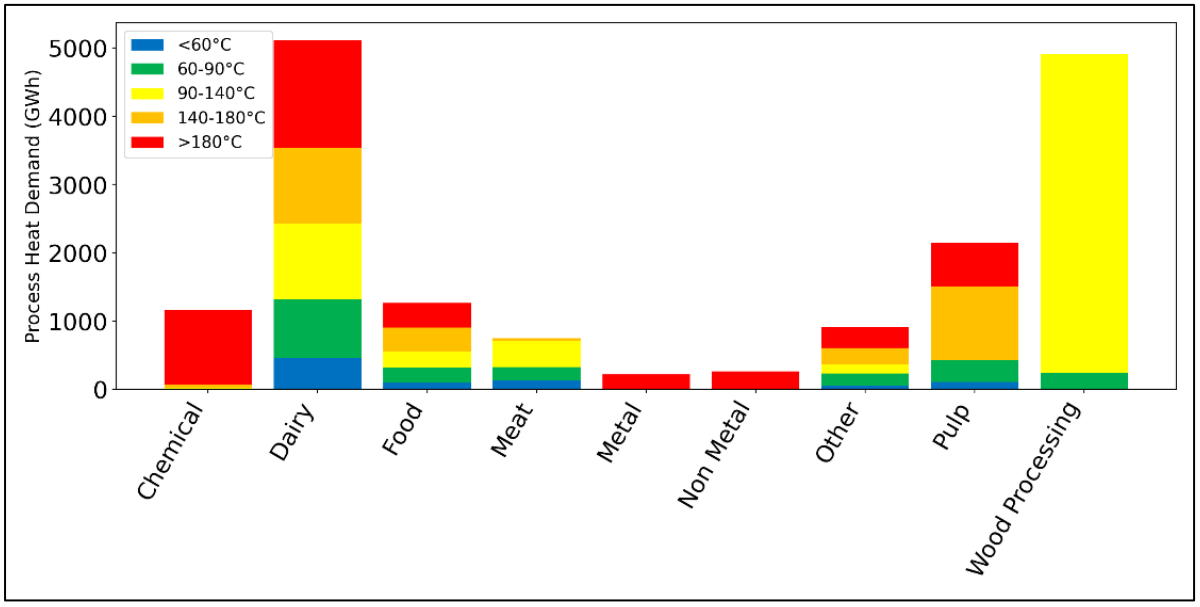


Figure 6.3: Sectoral distribution of decarbonisable process heat temperature levels in New Zealand, based on Table 4.2 and European Commission (2008), with additional referenced and partly confidential industrial data.

As EECA categorises heat temperature into three categories: high temperature (above 300˚C), intermediate temperature (100-300˚C) and low temperature (<100˚C), the standard framework is helpful to assess the quality of heat required for decarbonisation studies (EECA, 2024). As introduced in Section 4.4.1, the temperature is categorised based on heating technologies, consisting of a broad range of heating technologies, while the temperature shares are organised by process category.

Referring to Figure 6.3 and Figure 6.4, the cumulative heat energy demand curve shows that the top 24 highest energy-consuming sites account for 50% of the total decarbonisable process heat demand, while the top 68 sites account for 75%. In contrast, more than 100 smaller sites each contribute less than 1% of the total demand. To maintain the validity of the model, only major sites with decarbonisable heat demand greater than 5 GWh per annum are included. This approach is based on the assumption that factories with low heat demand have minimal impact on grid infrastructure changes and biomass consumption, while excluding them helps to reduce the computational burden of the model. Hence, there will be 298 factory sites considered in the model for decarbonisation of process heat. The model will include 99% of the total indirect allowable process heat for decarbonisation.

The temperature profiles in a factory site often generalised by the sector and processes within the plant due to the sheer number of the factory data. The data of end-use temperature is confidential; hence, it is adjusted with available literature to generalise the classification temperature demand profile for each process. Table 4.2 in Chapter 4 can be referred for evaluating the temperature bands.

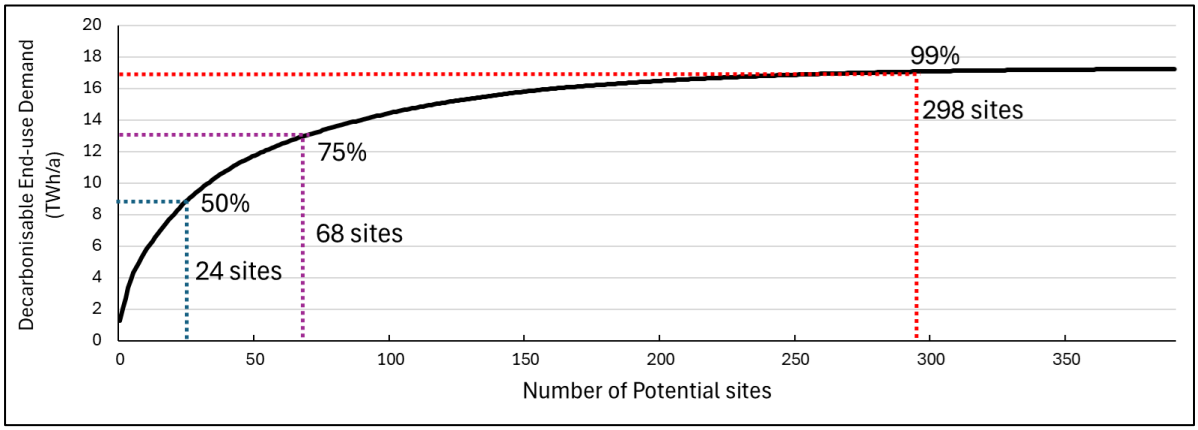


Figure 6.4: Decarbonisable process heat demand ranked in descending order with cumulative distribution.

In summary, from the compiled databases from the Heat Plant Database and the RETA database, there are a total of 429 sites. In the model, there are in total 298 major sites with more than 5 GWh per annum of decarbonisable process heat load. In total, there are 106 sites identified only from RETA, where only 130 sites are overlapping sites, and 62 only appear in the Heat Plant Database, as shown in Figure 6.5, which are being selected in the optimisation model. Due to the constraints imposed on decarbonisable temperature levels and the minimum threshold of decarbonisable demand on a site, a total of 131 sites is excluded from the model. These sites are

primarily small-scale dairy, food, and meat factories and are therefore omitted from the analysis, as shown in Table 6.2.

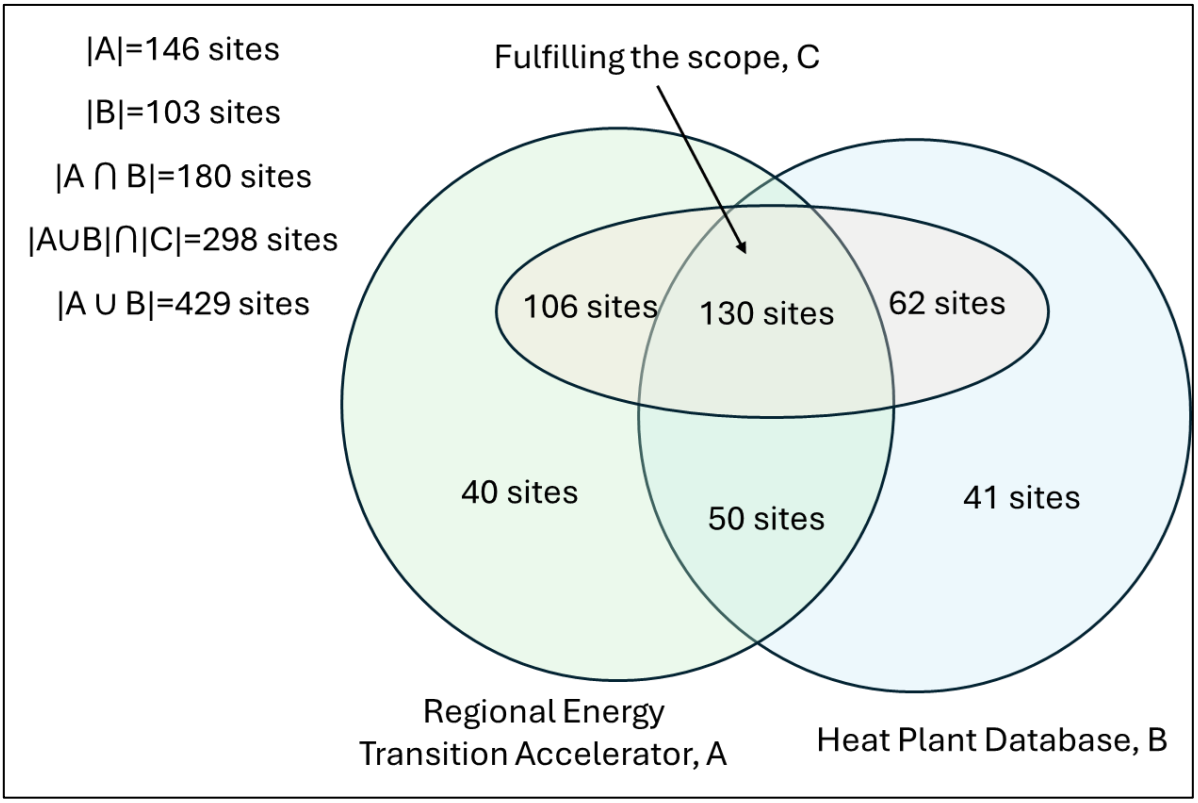


Figure 6.5: Compiled process heat dataset represented in a Venn diagram.

Table 6.2: Summary of industrial sites not included in the model and their associated process heat demand for decarbonisation (GWh yr^{-1}).

Sector	Both	HPD	RETA	Total
Chemical	5	3	9	17
Dairy	7	14	2	23
Food	7	11	8	26
Meat	21	7	8	36
Metal	1	2	2	5
Non-Metal	1	0	3	4
Other	1	0	7	8
Wood Processing	7	4	1	12
Total	50	41	40	131

For retrofitting coal boilers, there are multiple considerations when determining whether retrofitting is a viable financial and engineering decision within a plant. These considerations include cost compared to replacement, operational savings, equipment condition, adaptability of the boiler, logistics, and other factors. The energy systems model assumes that retrofitting is possible. It considers coal boilers with capacities greater than 5 MW that were installed after the 1990s to be eligible. This value represents a minimum viability threshold for retrofitting. Figure

6.6 and Figure 6.7 display the sites and boiler specifications where retrofitting is potentially feasible, along with the primary fuel used for process heat on-site.

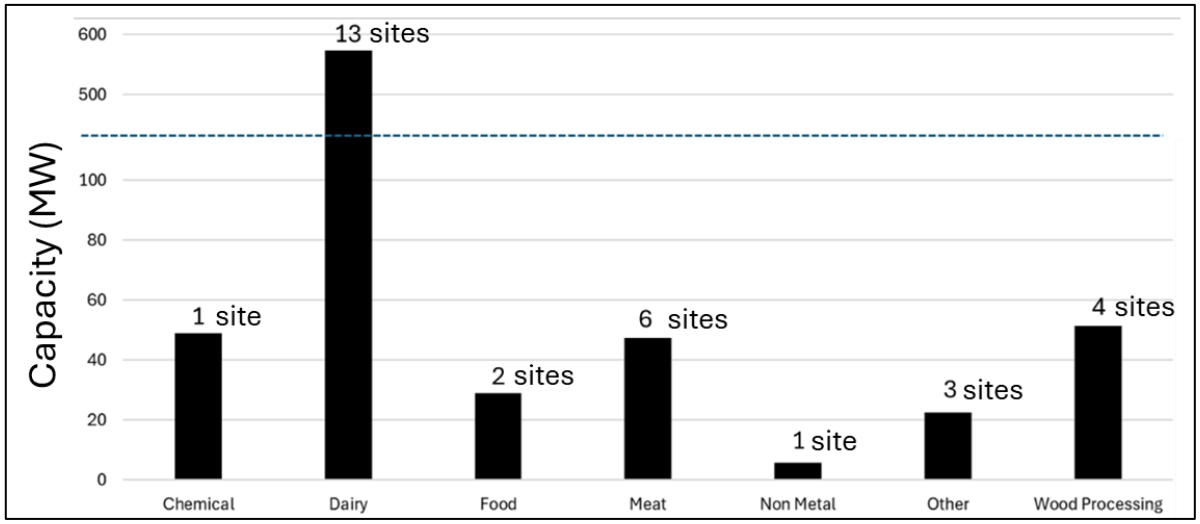


Figure 6.6: Number of coal boilers with retrofitting potential by industrial sector in the model.

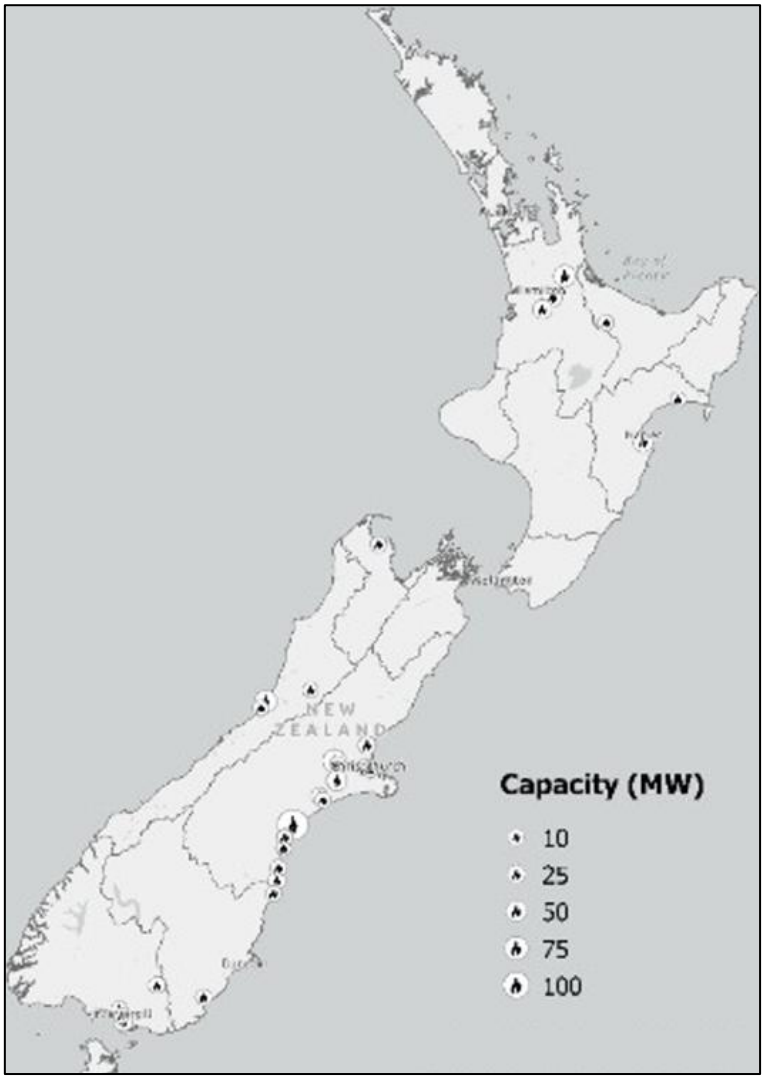


Figure 6.7: Spatial distribution of coal boilers with retrofitting potential by industrial sector in the model

6.4.4 Model Representation and Inputs

Industrial process heat demand is represented as discrete factory site nodes, each with an annual eligible heat demand, sector and process labels, current primary fuel, and geocoded location (Table 6.3). Each site is mapped to a region, island, and an assigned POC to link demand to the electricity system, while spatial location is used to calculate delivered biomass costs through distance-based transport parameters. Site heat demand is disaggregated into aligned temperature bands using sector and process coefficients from the literature, with temperature shares treated as a consistent input assumption where site-level profiles are unavailable. Finally, each site is assigned fixed eligibility rules, including coal boiler retrofit eligibility based on the specified threshold and technology availability constraints by temperature band and process class, which remain unchanged across scenarios unless varied in sensitivity analysis.

Table 6.3: Example industrial facilities used as model inputs for the nationwide process heat decarbonisation analysis.

Model Input	Example 1	Example 2
Object ID	15	53
Company Name	Fonterra - Te Awamutu	Alliance Group - Pukeuri
Industry	Dairy	Meat
Specific Process Type	Milk Powder	Meat
Energy Estimated for Decarbonisation (GWh yr^{-1})	183.67	13.93
Coordinates	-38.00, 175.31	-45.03, 171.03
POC Connected	TMU	OAM
Geographical Boundary (Island)	North	South
Region	Waikato	Otago
Current Primary Energy	Natural Gas	Coal
Social Aspects: (Urban/Rural)	Medium urban area	Rural other
Retrofitting Capability? (MW)	-	5.2
Current Boiler Year	-	1986

6.5 Electricity Generation

6.5.1 Overview of the electricity generation sector

There are in total 80 generation companies, 62 retailers and 6 gentailers (Electricity Authority, 2026) in New Zealand. The majority of electricity power generation is operated by gentailers, which are Contact Energy, Genesis Energy, Meridian Energy, Mercury Energy, Contact Energy and Nova Energy. The different technology types, such as onshore wind, utility solar, geothermal, co-generation, hydroelectricity, thermal, and bioenergy, are also compiled in the data collection step. Despite having proposals on tidal, marine, offshore wind, and pump hydroelectricity projects by generation companies, all these projects have been stalled or on hold due to economic considerations.

The electricity generation data is obtained from different sources. This is because the database includes an amalgamation of technologies but in some instances aggregates certain energy schemes into a single node of generation site mainly concerning hydroelectricity schemes. Hence, information on certain power stations can be obtained from the Electricity Authority, as well as from gentailer websites. For example, the Waikato Power Scheme is represented by eight sequential hydroelectric stations: Aratiatia (90 MW), Ohakuri (112 MW), Atiamuri (84 MW), Whakamaru (100 MW), Maraetai (360 MW), Waipapa (51 MW), Arapuni (175.2 MW), and Karapiro (108 MW) (Mercury NZ Limited, 2026).

6.5.2 Generation portfolio changes and capacity adequacy

The push for new renewable generation has become increasingly important as some thermal power plants have been decommissioned. For example, Contact Energy's 44 MW Te Rapa cogeneration plant, which had been operating since 1999, closed in 2023. Other closures include the 250 MW thermal Huntly Unit 3 in 2012 and the 404 MW thermal Otahuhu B in 2015, along with several other large-scale generation assets. The annual installed capacity has been displayed in Figure 6.8. It is seen that the uptake of solar generation has been increasing exponentially since 2011, whereas wind generation capacity has surged from 2014 until 2019. To ensure a continuous supply of electricity generation, natural gas is still operational and can generate up to 23% of the total electricity demand when solar and wind generation outputs are low.

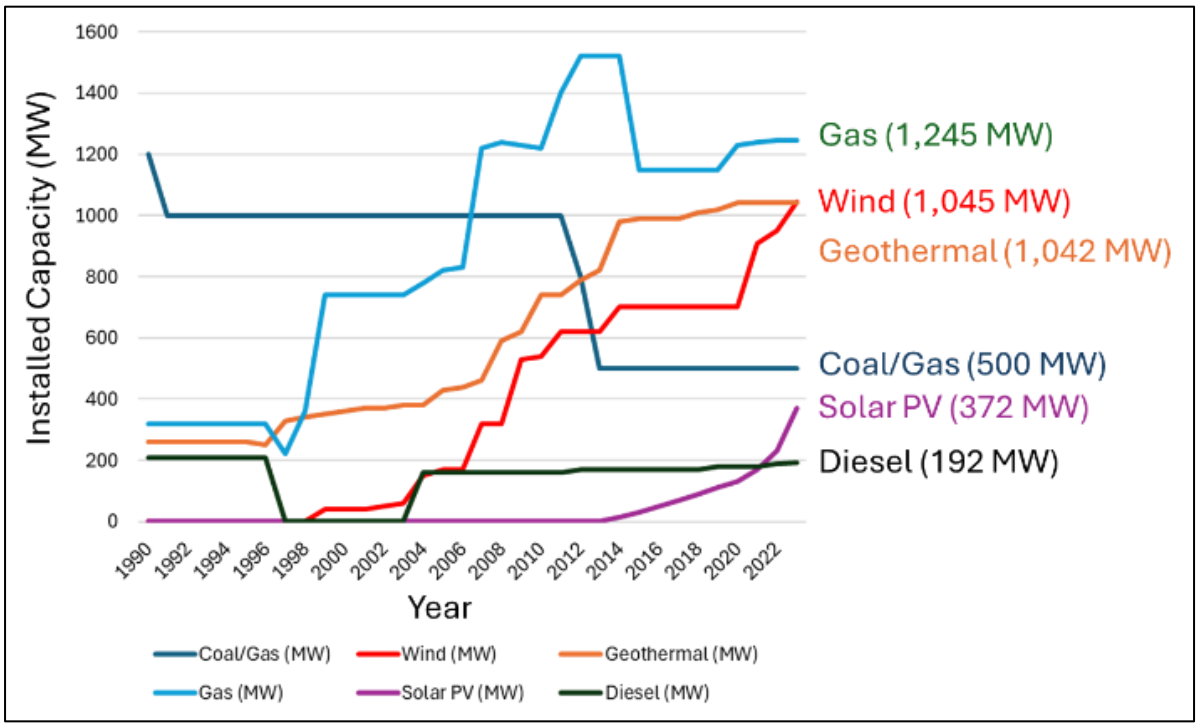


Figure 6.8: Installed capacities by plant type (excluding hydroelectricity) (MBIE, 2024b).

By including the total generation capacity with the daily peak demand and annual load data from 2010 to 2022, it is seen that there is an upward trend in peak demand which might cause outages when demand exceeds generation. It is seen also that the peak ratio and load ratio have to be consistent (Figure 6.9), especially when modelling the generation so as to prevent electricity

shortages during peak demand periods (Figure 6.10). Surplus operational capacity is incorporated into the model, ensuring that the necessary generation capacity costs are fully considered in the total system cost.

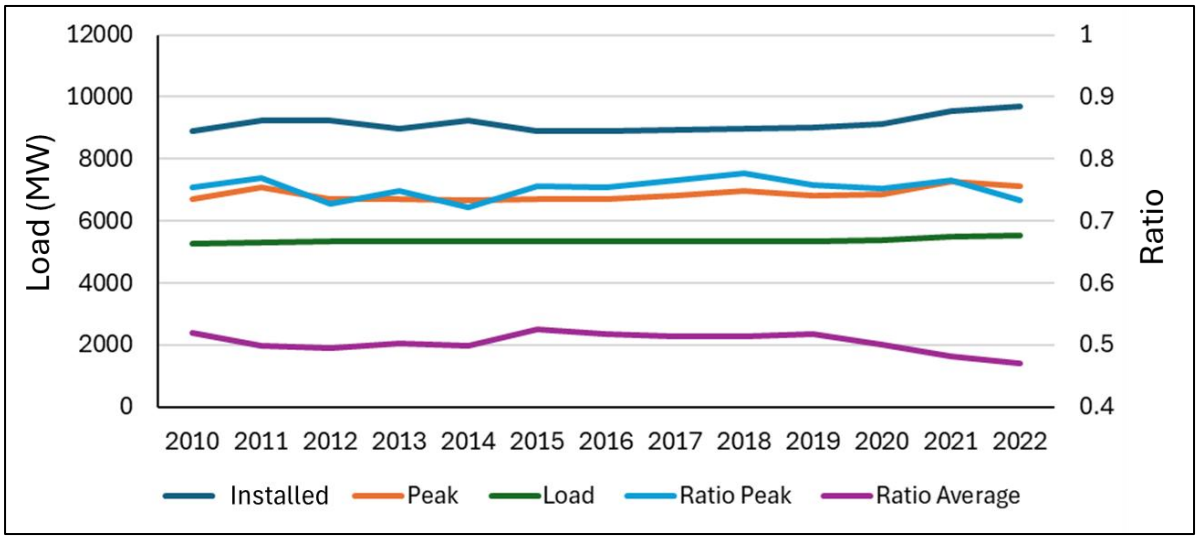


Figure 6.9: Comparison of peak demand ratio and load ratio from Jan 2010 to Sep 2022 (Transpower New Zealand, 2023b).

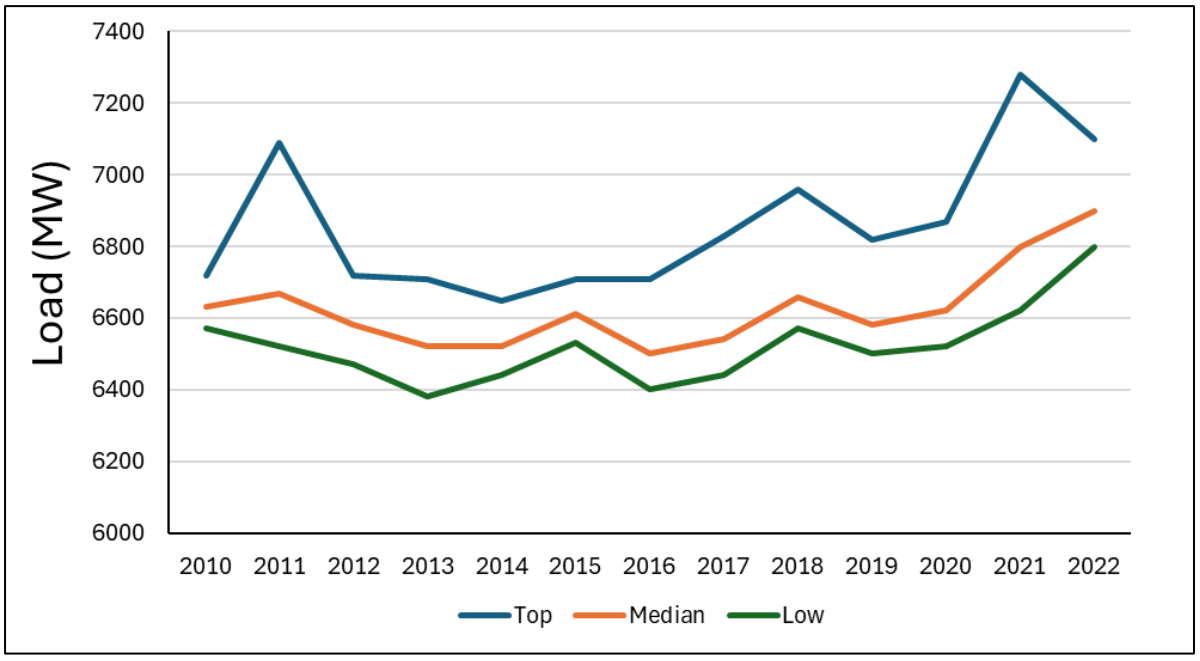


Figure 6.10: Top 20 daily peak electricity demand values from January 2010 to September 2022 (Transpower New Zealand, 2023b).

6.5.3 Preprocessing Tool

Electricity generation inputs require preprocessing where raw datasets are incomplete or do not provide model-ready time series. This section describes Renewables.ninja, which is used to generate consistent wind and solar generation profiles for all candidate sites.

Renewables.ninja

As all solar and wind generation are intermittent, the output is unpredictable, making operational generation not possible. They pose a challenge when trying to support the proposed case study model which requires time-series parameters and spatial differences between locations. As the solar and wind generations are also heavily location and weather dependent, they raise an urgency to develop a renewable power profile with high fidelity. Renewables.ninja is used to produce high temporal resolution for each renewable generation point, where the data is retrieved from the global Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) reanalysis model (Pfenninger and Staffell, 2016).

A key modification has been made to the tool by enabling aggregation and automation to obtain generation profiles for all studied sites consistently. In the New Zealand case study, 106 sites are included, comprising 54 wind sites (21 commissioning and 33 proposed) and 52 solar sites (9 proposed and 43 commissioning), including 13 utility-scale sites and 38 community-scale generation locations. As rooftop solar installations are located across residential and commercial buildings, they are widely distributed spatially. The spatial distribution of the intermittent generations, both commissioning and proposed are presented in Figure 6.11.

With the availability of Renewables.ninja source code in R, a Python-based API system, was developed and modified to allow automated requests for wind or photovoltaic generation data for every entry of location site. The workflow reads site coordinates from a spreadsheet and automates retrieval over extended periods of years, enabling efficient construction of the time series inputs required for the model. The full source code for the Renewables.ninja will be displayed in Appendix E.

As a summary for the data preparation tools, Renewables.ninja is helpful for generating datasets, which are an important parameter for the input of the energy model.

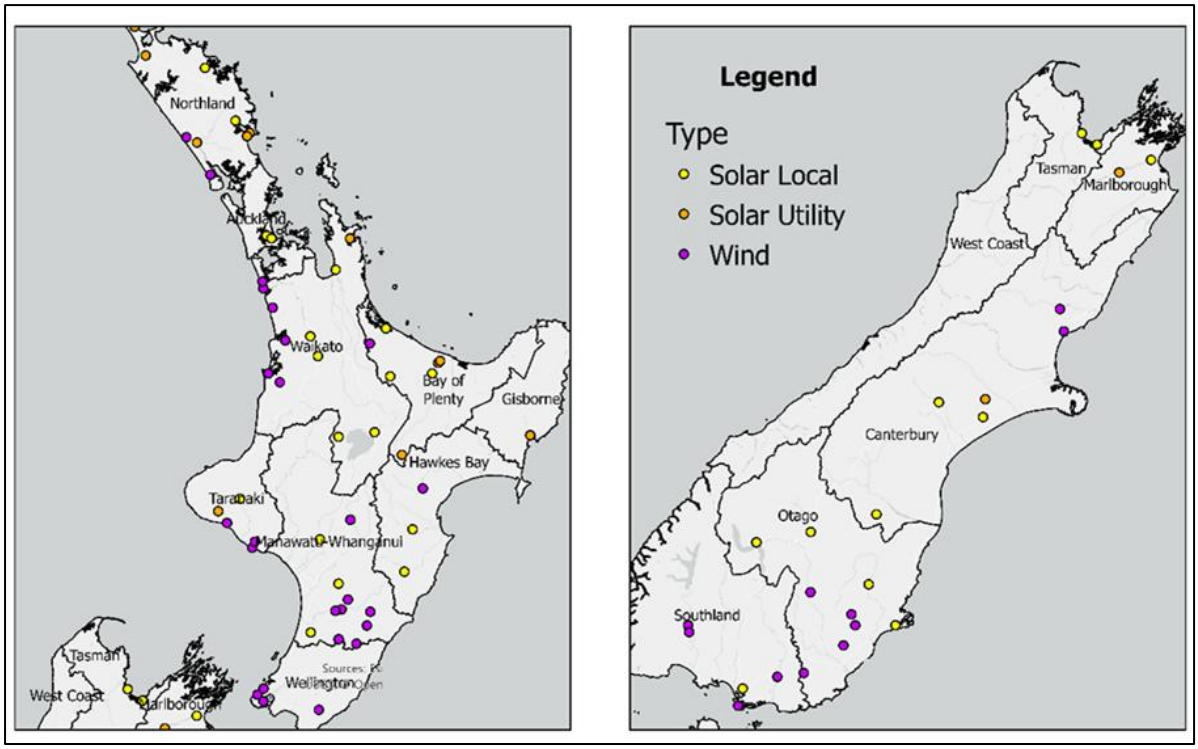


Figure 6.11: Spatial distribution of intermittent generations in New Zealand.

6.5.4 Model Representation and Input

Electricity generation is represented as discrete generation site nodes, each associated with generation technology type, capacity, operational status, operator, region, island, and POC. As shown in Table 6.4 these attributes support spatial linking of generation to the electricity network representation and enable technology-specific cost and performance inputs to be applied within the model.

Table 6.4: Example input parameters for electricity generation sites in the model.

Specifications	**Example 1**	**Example 2**
Object ID	9	401
Types of Power Station	Solar	Hydroelectric
Type	Local	-
Generation Name	Central Canterbury	Maraetai I
Year	-	1954
Region	Central Canterbury	Waikato
Capacity (MW)	35.7	252
Operator	Orion New Zealand	Mercury Energy
Status	Commissioning	Commissioning
POC Connected	COL	MTI
References	Electricity Authority (2025)	Mercury NZ Limited (2026)
Geographical Boundary (Island)	South	North

6.6 Biomass Resource

6.6.1 Biomass Overview

Biomass fuels such as in-forest residues, sawmill chips, k-log wood, straw and stover are renewable energy that can be replenished from the New Zealand forests. Biomass fuels are considered a renewable alternative to conventional fossil fuels. This is because biomass fuels can be replenished in a short period of time (Khare *et al.*, 2022). However, biomass fuels are only sustainable when the biomass is recovered from commercial forests. Figure 6.12 shows the spread of planted forests where the biomass can potentially be supplied for commercial uses, including heat generation uses.

In the model, the five types of biomass fuel shown in Figure 6.12 are selected as a potential source of process heat in industries. Wood pellets are manufactured from compressed biomass feedstock and compacted at high pressure, which increases their energy density for better fuel efficiency.

As local biomass resources are bounded by the country's resource constraints, additional use of wood pellets is unbounded and supplied by four major wood pellet retailers, which are Nature's Flame in Waikato, Azwood in Tasman, Niagara Sawmill and Timber in Southland and Otago Pellet Fires Ltd in Otago. Table 6.5 also outlines the specification of each biomass resource as its respective fuel type and its net caloric value for combustion. The overall cost of biomass comprises the cost of biomass feedstock, harvesting, collection and chipping cost. The North Island and South Island have distinct biomass costs, reflecting actual regional situations. The transport cost is assumed to be 0.30 NZD t^{-1} km^{-1}, which is uniform to all biomass types (Ministry of Transport NZ, 2024). However, biomass transport cost would be different, as they have different bulk density and net calorific value. Table 6.5 shows the cost differences and specifications between biomass types and locations. The price for each resource is obtained from Regional Energy Transition Accelerator reports (EECA, 2025c), and the net calorific value (NCV) of biomass is obtained from Hall (2024).

Table 6.5: Cost specifications by regions and net calorific value.

Biomass Type	**Price (NZD t^{-1})**		**Net Calorific Value (NCV) (GJ t^{-1})**
	North Island	South Island	
K-log wood	113	123	6.9
Sawmill chips	108	108	6.9
In-forest residues	70.7	76	11
Straw and stover	75	75	13.4
Wood pellets	484	484	17.5

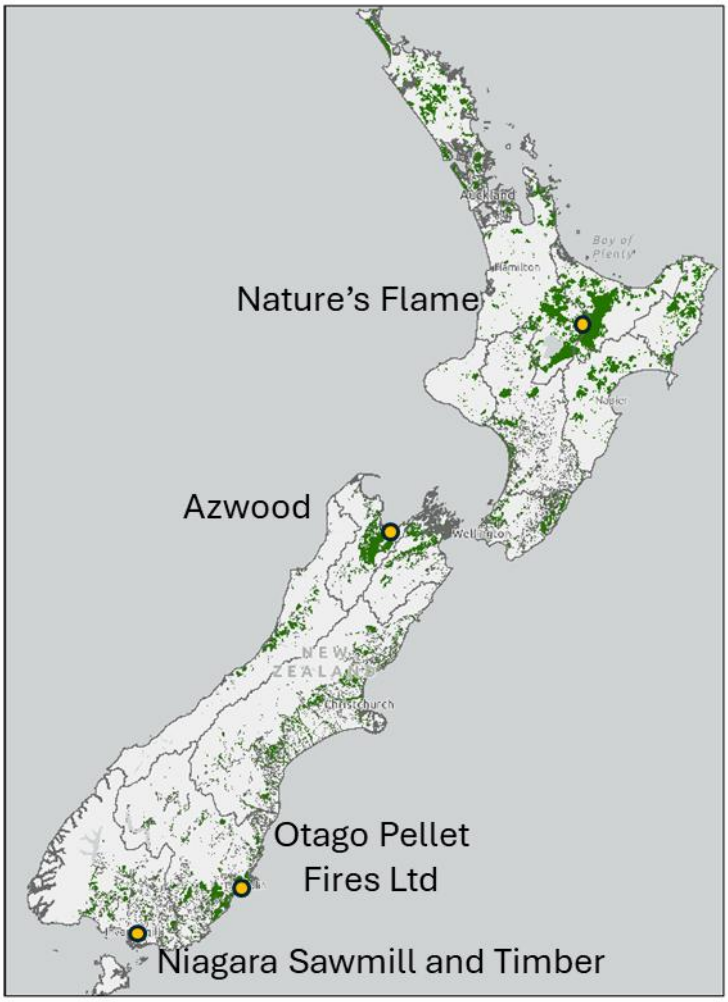


Figure 6.12: Agricultural and planted forest cover data from the New Zealand Land Use Map (Ministry for the Environment NZ, 2025) and locations of wood pellet suppliers.

Referring to Figure 6.13 and Figure 6.14, the Central North Island region, namely, Bay of Plenty, Gisborne, Hawkes Bay and Waikato, have a large volume of K-grade logs. In-forest harvest residues are typically abundant at locations with planted forests. Locations which are more industrial and commercially oriented regions, such as Auckland, Wellington and Taranaki region have the lower share of biomass resources due to the limited availability of agricultural land. Sawmill chips are found at locations with wood processing industries such as Bay of Plenty, Waikato and the Hawkes Bay region. The South Island has the largest share of straw and stover, with Canterbury and Otago dominating the share, amounting to 437,108 tonnes and 43,361 tonnes respectively, whereas Manawatu-Whanganui and Waikato are coming third and fourth with 31,798.4 tonnes and 31,016.36 tonnes respectively.

As these values represent gross volumes, a recoverability factor is typically assumed in the scenarios, since it is not sustainable to utilise all bioenergy available within forests. Instead, the analysis considers the proportion of biomass that can be sustainably recovered while allowing the resource to replenish over time. Table 6.6 presents the ranges of recoverability determined for

forestry resources. For the base case, the midpoint of the recoverability range is selected as the representative assumption.

Table 6.6: Recoverability factors based on different estimations by Hall (2022).

Biomass Type	L0 (Selected)	L1	L2
In-forest harvest	0.725	0.8	0.65
K logs	0.875	0.95	0.8
Sawmill Chip	0.65	0.7	0.6
Straw and Stover	0.625	0.75	0.5
Pellets	1	1	1

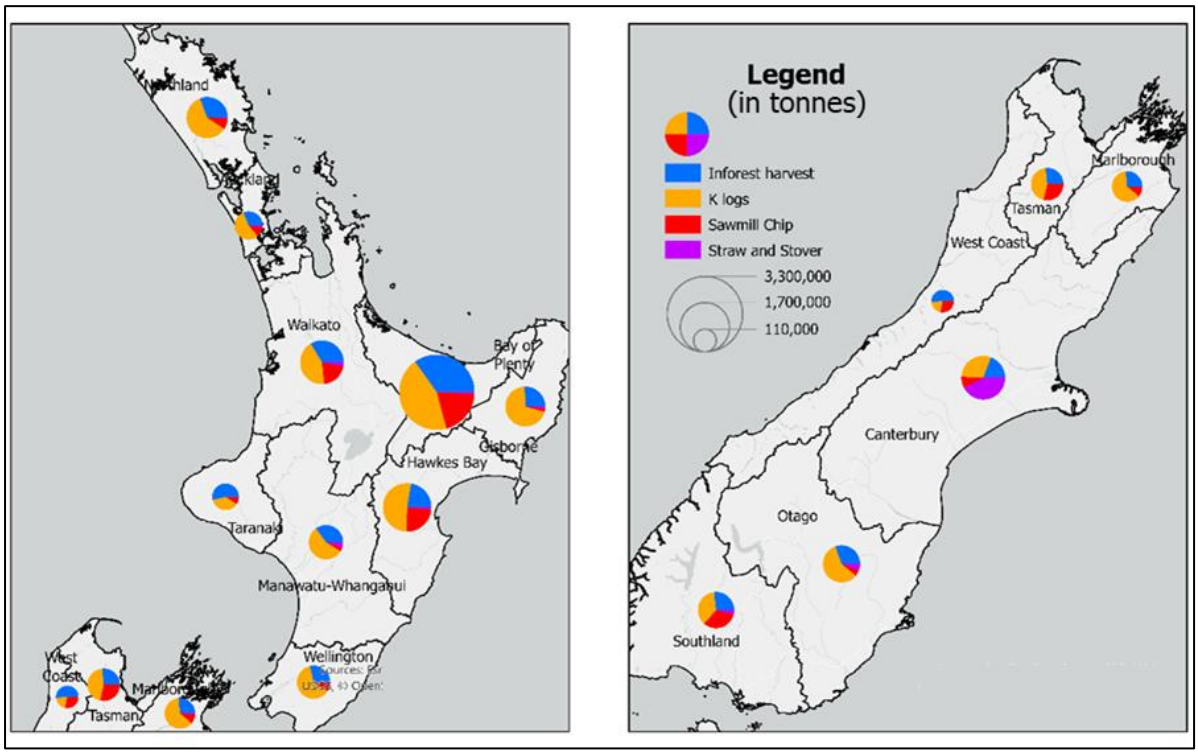


Figure 6.13: Spatial distribution of biomass resource types by gross volume (Hall, 2022).

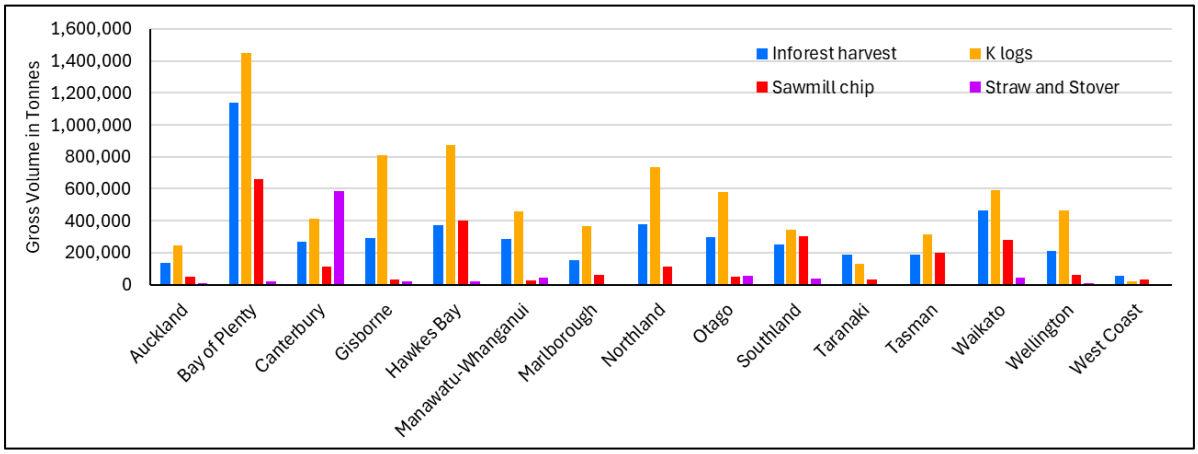


Figure 6.14: Bar chart on biomass resource type by gross volume by region (Hall, 2022).

6.6.2 Geospatial Tool

The biomass layer was developed in ArcGIS Pro through a series of spatial data processing and analysing steps before estimating the distance between factories and the biomass sites. The raw biomass layers representing each region were first extracted from a geospatial data repository. The feature within the layer for biomass is usually a polygon, as it shows the area coverage of the forestry resource of the region.

The features within each layer can be represented in various formats, such as points, lines, polygons, or rasters. A point refers to a single location in a map, defined by X and Y coordinates. A polygon represents a closed area on a map, which is constituted by a set of coordinates that form a boundary. A line in GIS represents a one-dimensional data which are joined together with a starting and end point. A raster is a dataset which arranges itself as a matrix of rows and columns containing a discrete or continuous value based on an attribute.

The biomass layer is extracted from the Ministry of the Environment New Zealand database (Ministry for the Environment NZ, 2025). The attribute tables of the file were first examined. As the database contains multiple entries such as natural forest, planted forest, cropland, grassland, wetlands, settlements and others, the unwanted biomass entries have to be isolated and removed from the layer. The remaining entries are further filtered by size to be at least 10 hectares so that any small polygon depicting biomass is eliminated to focus significant forest areas. This is followed using a tool from ArcGIS that is Multipart into Singlepart so as to disconnect combined pieces into different entries, which helps in merging a few biomass areas by location. As some forest areas in particular, the Waikato Region covers a large amount of area in a single entry, the area is manually partitioned into smaller areas to reflect realistic logistical resource collection. Subsequently, the entries are manually merged to ensure logical spatial divisions.

The Calculate Geometry tool was used to determine the covered area of each entry where the available biomass resource is proportional to the size of the area. Centroid points were generated from each polygon entry to represent the geometric centre of the corresponding forest area. As the centroid points do not inherently inherit attributes such as region and area, the Spatial Join tool was performed with the Regional Council layer from StatsNZ to identify the administrative region of each centroid (Ministry for the Environment NZ, 2025). The general workflow for the preparation on the biomass layer is as follows:

1. Data Inspection and Filtering – [Select by Attributes tool]
2. Declustering polygons with a single entry to independent entries [Multipart to Singlepart tool]
3. Partition large areas with single-entry [Split tool]
4. Aggregation of Areas – [Buffer and Merge tool]
5. Area calculation and extract coordinates [Calculate Geometry tool]
6. Centroid Points [Feature to Point tool]
7. Attribute Extraction [Spatial Join tool]

The final location of biomass nodes in the New Zealand case study is shown in Figure 6.15.

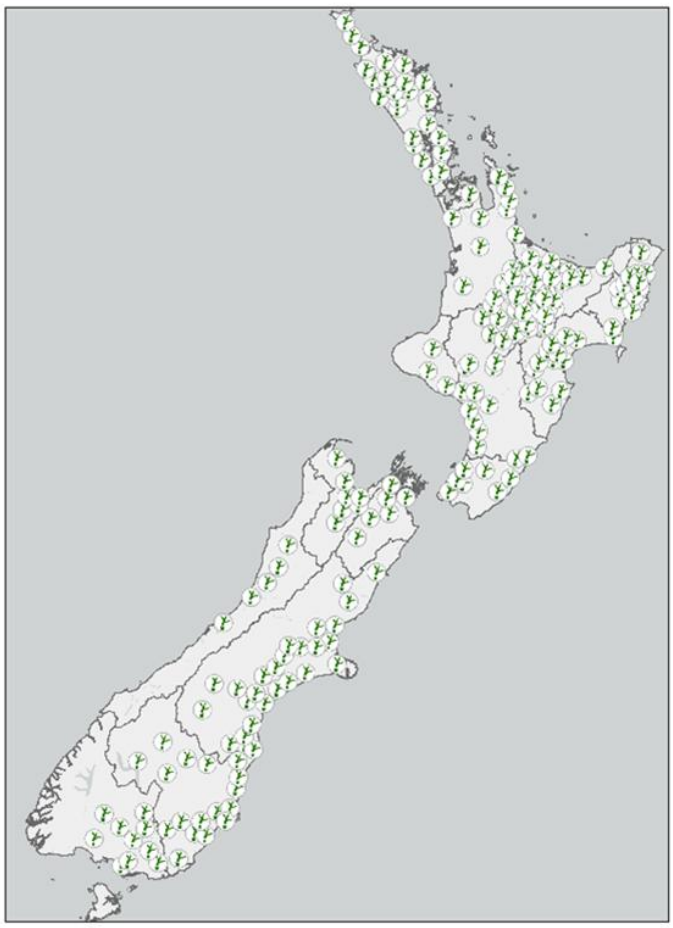


Figure 6.15: Biomass location nodes in New Zealand, identified using geospatial analysis.

6.6.3 Model Representation and Input

Biomass resources are represented as discrete nodes, each associated with biomass type, resource limit, spatial attributes, region, island, and forest cover. These attributes support spatial linking of biomass to the industrial process heat sites for biomass-based boilers and thermal generation sites for retrofitting gas turbines, as shown in Table 6.7.

Table 6.7: Example biomass center inputs.

Specifications	Example 1	Example 2
Biomass ID	9	135
Forest Area (km^2)	439.279	119.347
Region	Otago	Waikato
Coordinates	170.048, -46.081	176.101, -38.173
Geographical Boundary (Island)	South	North
In-forest harvest (gross tonne)	50,276.0	10,081.9
K-log wood (gross tonne)	115,043.1	15,345.6
Sawmill Chip (gross tonne)	7,363.2	5,322.3
Straw and Stover (gross tonne)	9,084.5	845.2
Wood Pellets (gross tonne)	-	-

6.7 Point of Connections (POCs) and Electricity Grid Assets

6.7.1 Market structure and operational context

New Zealand's electricity market is a decentralised electricity market which is regulated by Electricity Authority. The initial intention for the design of the wholesale electricity market was to not set a fixed price cap or a no capacity market, with the hope of encouraging competition in a deregulated market (Philpott *et al.*, 2019). Each trading period is at 30-minute intervals. Electricity generators offer supply to meet demand, with reserve peaking thermal power plants available to respond to sudden faults or generation shortfalls.

Generators submit bids for each trading period, which are ranked from lowest to highest in a price stack for dispatch. In each trading period, generation is selected from the cheapest to the most expensive until demand is met. Then, all dispatched generators are then paid the same price, corresponding to the highest accepted bid in the stack, known as the spot price for that trading period. In principle, nodal pricing supports supply–demand balancing across regions by allowing regional prices to differ according to the market (Poletti, 2021). Figure 6.16 shows how the spot electricity market works in New Zealand.

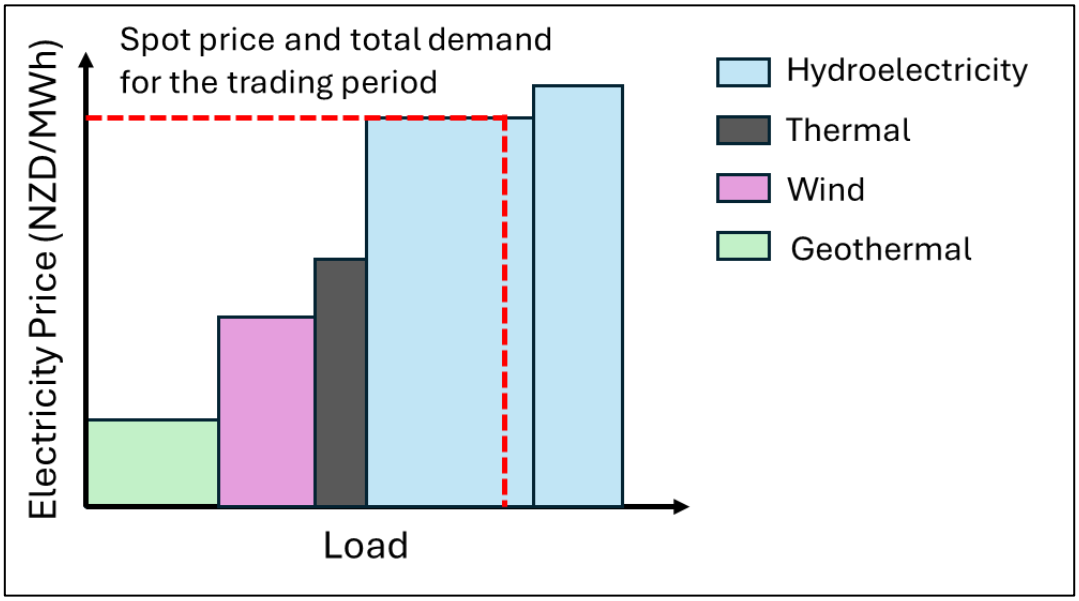


Figure 6.16: The concept of spot electricity market pricing in New Zealand.

This pricing structure allows electricity to remain relatively low-cost during periods of low demand. However, it can also lead to high prices during system faults or peak demand periods. Figure 6.17 shows the historical real-time spot electricity prices in the New Zealand market. This structure may also affect affordability, as generators are not strongly incentivised to invest in new capacity due to marginal pricing, which requires them to recover investment costs (Poletti, 2021). This places financial risk on generators, particularly for intermittent technologies such as wind and solar, which are less predictable and have limited dispatchability (McRae, 2025). Moreover, real-time price volatility can be unfavourable for industrial users, as it may disrupt process operations or to bear the high electricity costs.

In summary, most consumers, particularly large energy users, mitigate price volatility through hedge contracts such as fixed-price variable-volume (FPVV) contracts, where electricity prices are negotiated and locked in to reduce exposure to spot market fluctuations.

Energy security has also been an issue, as the ban on offshore exploration for natural gas was announced in 2018, and this increased the number of dry years. Based on historical records shown in Figure 6.17, 2015 was a dry year, with an average spot price hovering around NZD 140 MWh^{-1} and prices spiked up to NZD 4,000 MWh^{-1} on the 2nd June 2016 when the lake levels were below the acceptable threshold. Electricity prices also peaked at NZD 893 MWh^{-1} in August 2016 due to low hydro inflows Figure 6.17 displays the historical spot electricity price by trading periods.

As a summary, most consumers especially large energy user would take out hedge contracts namely fixed-price variable-volume contracts where the electricity price is negotiated and locked-in to prevent the need to face electricity fluctuation price.

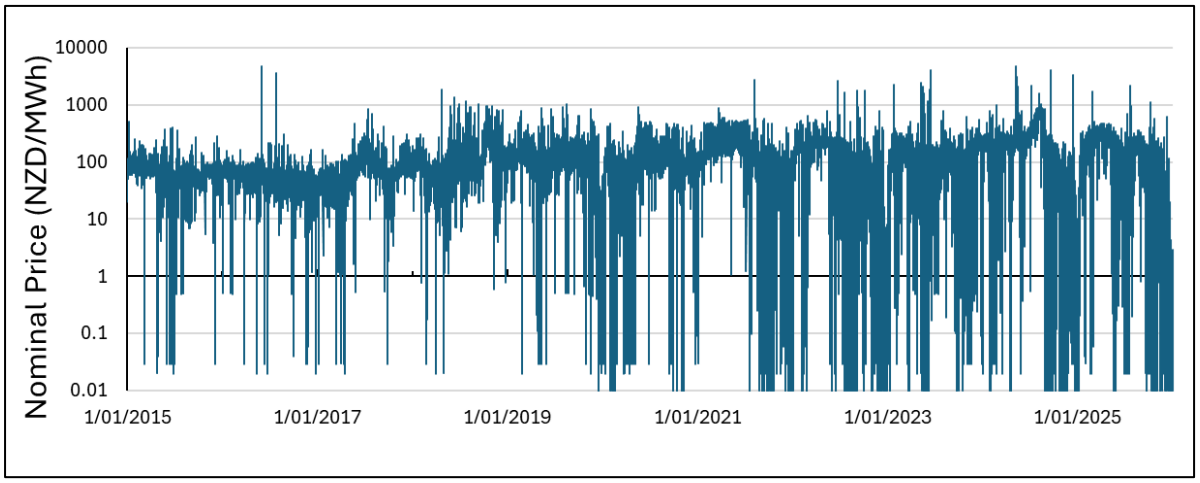


Figure 6.17: Wholesale electricity prices (logarithmic-scaled) from 2021 to 2025 at hourly resolution, based on data from the Electricity Authority (2025).

6.7.2 Estimations on Residential, Commercial and BEV Demand

Referring to Figure 6.18, residential electricity use occupies 13.74 TWh of the electricity is a sizable amount as people commute to work (MBIE, 2024b). Additionally, commercial electricity use is 9.52 TWh whereas industrial electricity is 13.72 TWh. Battery electric vehicle (BEV) which is currently at 231.6 GWh per annum in 2022 and is projected to increase to 2,692.1 GWh in 2040, an increase by tenfold in electricity consumption (MBIE, 2024b).

However, due to the spatial nature of the model, it is not sufficient to rely only on national sector totals. The model requires demand to be represented at the POC-level. Each POC is assigned an FPVV electricity price, which is formulated using wholesale electricity price data from the Electricity Authority. At the same time, electricity demand for each trading interval is recorded by the Electricity Authority and can be downloaded and used as model inputs. Because the datasets contain very large numbers of entries, preprocessing tools are required to extract the data reliably.

One of the main barriers for public websites is that data extraction is limited by download size and by the number of entries that can be retrieved at once. In the New Zealand case study, the relevant high-temporal-resolution datasets are wholesale electricity prices and electricity demand for each trading period (30-minute intervals). In the Electricity Authority database, there are 242 GXP points with alternating current (AC) stations where wholesale prices are disclosed. To increase sample size for statistical analysis, including the Monte Carlo analysis used in Chapter 9, three years of electricity data are extracted through the portal.

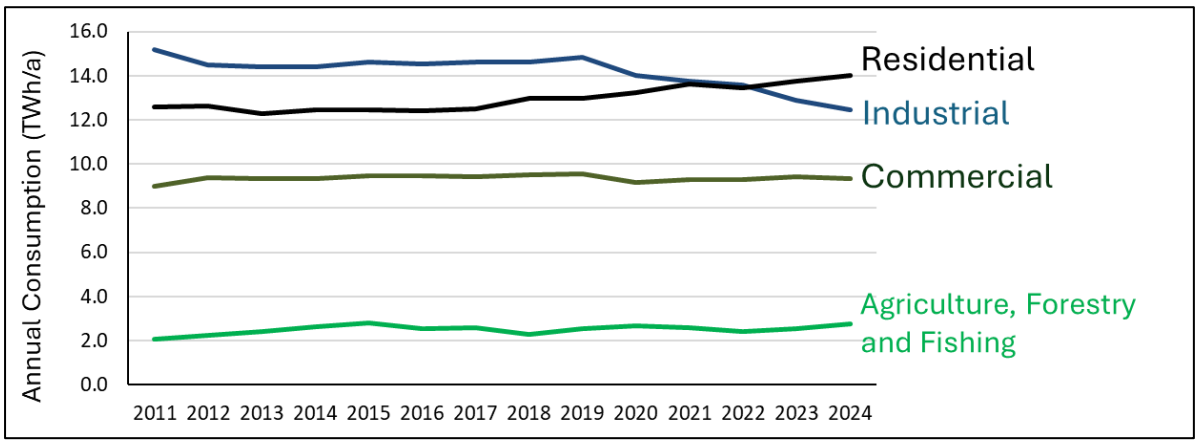


Figure 6.18: Electricity consumption by sector based on data from MBIE (2024b).

6.7.3 Assumptions on the Electricity Price (FPVV)

There are several pricing methods for electricity supply for businesses in New Zealand. The typical pricing methods are FPVV, proportional pricing and spot market pricing. The FPVV method refers to having a negotiation with the energy company to offer a fixed energy rate. The method ensures that the electricity price is predictable and stable, and the change in electricity price for every trading period, which is a 30-minute interval, will not affect the business operation.

The second method is the spot market price. Companies using the spot market price scheme would be affected by the wholesale electricity prices at every trading period, often suited for plants which are not heavily dependent on electricity use and possess the ability to ramp up and ramp down operations with the change of the electricity price. In short, companies will have to pay the price set by the wholesale electricity market along with additional charges, including retailer operating costs, metering costs, network charges (such as transmission and distribution) and levies. The third method combines both methods, where a percentage of the electricity is priced at a FPVV rate, while the remaining is billed at the spot market rate.

In the energy model, the electricity charges for industrial companies are assumed to be FPVV rate. This is because most large energy users typically opt for FPVV rate with contracts of 21.9 TWh in 2023 and 22.4 TWh in 2024 (Hensen, 2025). The weighted FPVV is often confidential, which is between the price negotiation between the energy user and the retailers for the volume of purchase. The values used in the model were also adjusted based on the average annual spot

electricity price, indicating that it is a benchmark for a proxy for varying electricity prices for several years.

The FPVV cost formulation used in this energy model is developed by the Ahuora Energy Centre, where the detail of the weighted coefficient is confidential. However, the general overview of the electricity cost trend is separated into month-based, day-of-the-week, hourly-periods-based and region based. As a general statement, the price ratio is higher in winter months, especially in July and August, whereas electricity price would be lowered considerably in January and December. Electricity price is generally lower during the weekend compared to weekday where there is high demand due to work productivity.

The Electricity Authority database provides information on POC demand and wholesale electricity prices. The demand is added up to monthly totals, and the wholesale prices are processed further using Ahuora's FPVV calculator. It is important to note that the sales of electricity price applied at factory inflow constitute the generation and network charges rather than just the generation itself (Electricity Offers NZ, 2025). Hence, it not only includes the cost of generation but also a share such as network cost, which includes demand charges and daily fixed charges as shown in Table 6.8. Due to the lack of availability and transparency which reports the electricity breakdown of individual sales costs for large energy users, as most prices are set based on negotiation with the energy companies, a cost assumption has to be made.

The applied electricity price represents an estimated sales tariff constructed using FPVV for the variable component and an assumed allocation for demand and fixed charges. The FPVV component is assumed to constitute 60% of the total sales tariff, reflecting the upper bound of reported values for large industrial consumers, whereas network charges and other miscellaneous charges account for 40% of the sales of electricity.

Table 6.8: Estimated cost components fractions for large energy users (Electricity Offers NZ, 2025).

Cost Components	**Price Range for the Total Electricity Bill**
Unit Rate (NZD kWh^{-1})	40-60%
Demand Charges (NZD kW^{-1})	20-35%
Daily Fixed Charges	5-15%
Network Costs	Included in above

Figure 6.19 shows the aggregated relative spot price ratio across all GXPs in New Zealand for the period 2022 to 2024. Although spot prices are different across trading periods and GXP locations, the relative price ratio is not always higher or lower for any GXP. For modelling purposes, the relative price is adjusted and normalised across the defined time blocks to reflect spatial variation while maintaining consistency within the optimisation framework. Therefore, the factory electricity price $C_{l,t}^{FPVV}$ is expressed as:

$$\begin{aligned} Factory\ Electricity\ Price, C_{l,t}^{FPVV} \\ = Spot\ price \times GXP\ ratio \times FPVV\ formulation \\ \times Additional\ Component\ Charges \end{aligned} \quad (6.1)$$

Table 6.9 presents the electricity price ranges for the key GXPs represented in the model, after accounting for spot electricity prices, GXP-specific price ratios, the FPVV formulation developed by the Ahuora Energy Centre, and additional components charges.

Table 6.9: Electricity price ranges for key GXPs implemented in the model.

GXPs	**Benmore (BEN)**	**Otahuhu (OTA)**	**Islington (ISL)**	**Stratford (SFD)**
Price (NZD MWh^{-1})	98.8 – 238.6	117.9 – 270.3	104.1 – 258.3	110.3 – 242.4

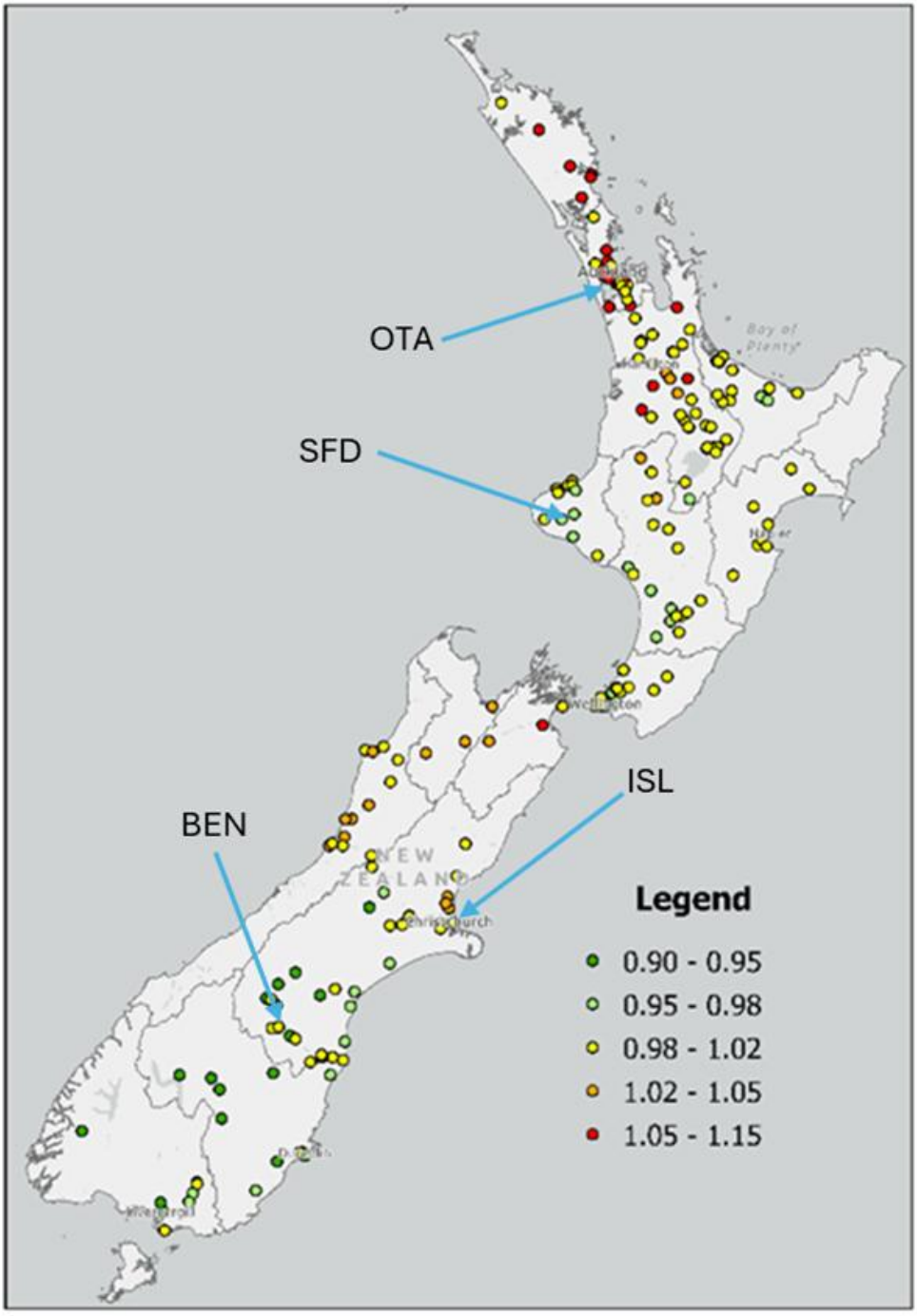


Figure 6.19: Adjusted price ratio across GXPs.

6.7.4 Grid Connection and Security Modelling

POCs that exclude TEE junctions and HVDC stations are usually grid injection point (GIP) nodes and grid exit point (GXP) nodes. The National Grid comprises 52 GIP nodes and 196 GXP nodes (Transpower New Zealand, 2022a). Fundamentally, most GIP nodes are also GXP stations; POCs such as Islington (ISL), Arapuni (ARI), Huntly (HLY) exhibit both capability of allowing electricity to enter and leave the grid. Conversely, POCs such as Penrose (PEN) and Hamilton (HAM) are solely GXPs, and POCs such as Rangipo (RPO) is the only station which only introduces electricity to the grid (Ergo Consulting, 2024).

In Chapter 7, POCs are used to spatially allocate electricity demand and generation to nodes, but transmission constraints and n-security are not enforced because the electricity system is represented as a copper plate. Grid constraints will be revisited to demonstrate the importance of the calculation of grid assets in Chapter 8. This is to show the difference between single deterministic optimisation and multi-stage deterministic optimisation.

N-security Demand and Generation

For the New Zealand case study, N-security demand limits are obtained from the Transpower Planning report (Transpower New Zealand, 2023a) and the ERGO Consulting report on RETA (Ergo Consulting, 2024). N security generation limits are extracted from the Transpower Envision spatial tool (Transpower New Zealand, 2025a). These values are input constraints to represent grid security margins and prompt for upgrades at the POC level.

Challenges arise in implementing full N-1 security. It is not about the modelling limitations, but it is the availability of data. Detailed contingency level information is required to model N-1 conditions, such as the specific power outages of transmission assets and the dispatching limits for every major node, if they are not publicly available. Due to the shortcomings on the data, the case study applies the N security limits as a proxy for secured grid constraint limitations while also highlighting that the grid model does not reproduce a complete N-1 contingency grid configuration.

Demand based on Sectors

National population projections from Stats NZ are used to support demand assumptions. The reference year is set to 2022. Due to differences in climate and temperature by region, electricity use varies spatially. In particular, northern regions consume less electricity for heating, whereas electricity consumption per household increases at lower latitudes. Figure 6.20 shows average consumption using a colour scale for different regions based on the approach made in Chapter 4.4.3.

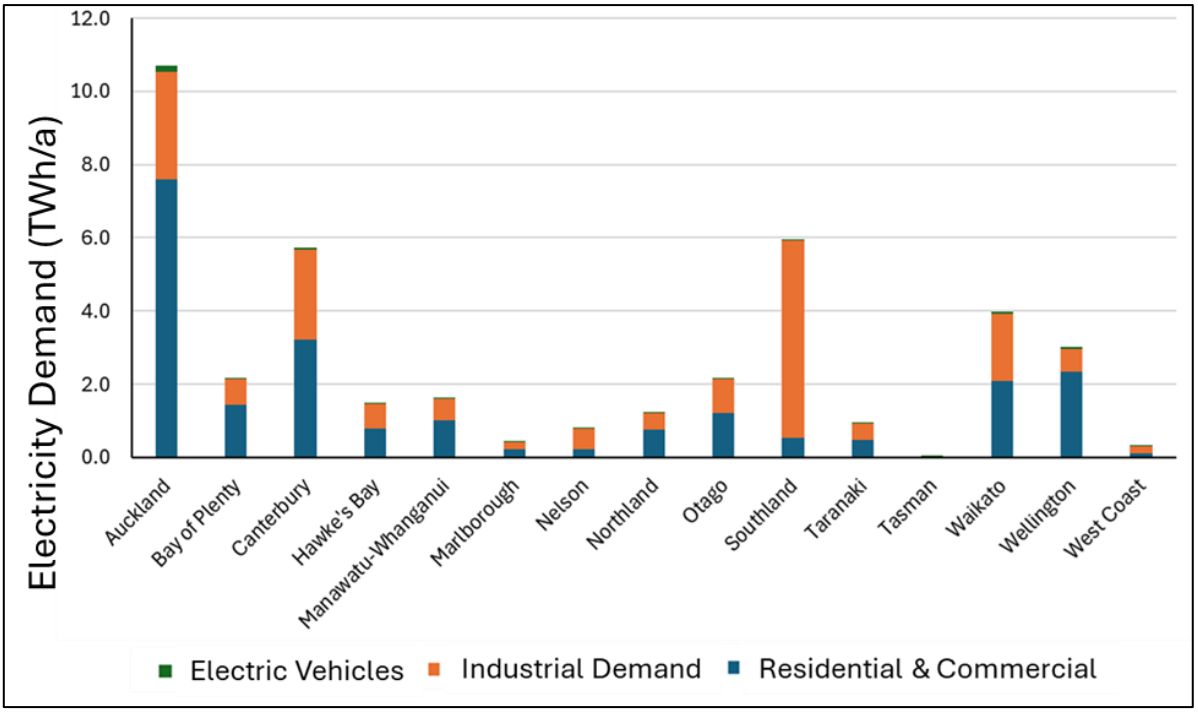


Figure 6.20: Electricity consumption in 2022 by sector (EV, industrial, residential, and commercial), based on model assumptions.

Transmission Grid Lines as a System Constraint

New Zealand's national electric power transmission system is known as the National Grid Transpower New Zealand, a state-owned company, manages, governs, and maintains the National Grid. Transpower operates more than 11,000 km of transmission lines across the country, supported by approximately 217 POCs (Transpower New Zealand, 2025b). Out of 217 POCs, there are 29 TEE junctions which join POCs together and also function to branch off from a larger voltage transmission line to supply electricity to a local distribution network. There are also 6 HVDC stations in the transmission grid, which is part of the HVDC inter-island link. The HVDC stations have an installation of 1200 MW bipolar capacity which transfers electricity bi-directionally between the North and South Islands to match the energy availability and demand in the country.

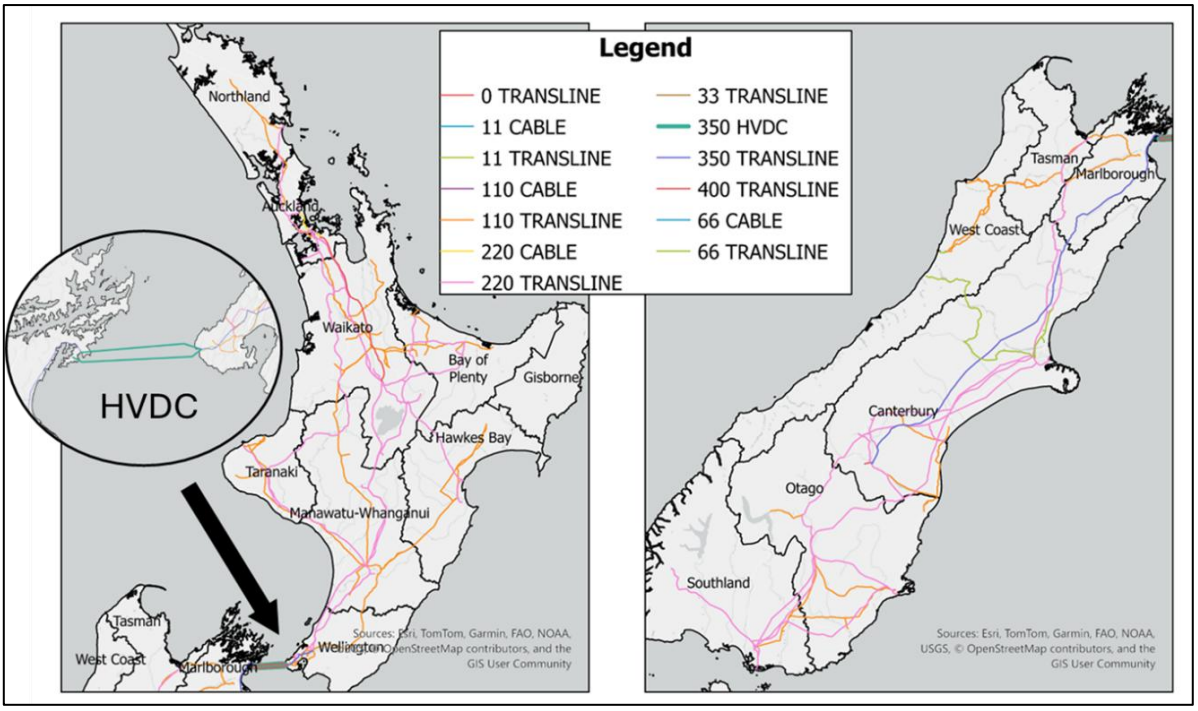


Figure 6.21: Spatial representation of POC and transmission layers, including HVDC infrastructure, across New Zealand with data from Transpower New Zealand (2025b).

6.7.5 Model Representation and Input

POCs are represented as discrete nodes with attributes describing location, region, island, station type, electricity demand, process heat potential, commissioning generation, and grid security parameters in Table 6.10. Wholesale electricity prices are represented at the POC level for the reference year.

Table 6.10: Example POC input specifications, including technical and demand-related parameters for selected nodes.

Specifications	Example 1	Example 2
Object ID	6	203
POC	ASB	MKT
Full Name	Ashburton	McKee Tee
Geographical Boundary	South	North
Region	Canterbury	Taranaki
Type	AC Station	TEE
NZTM-x	1503871	1711710
NZTM-y	5133909	5678043
Electricity Demand (GWh yr^{-1})	665	-
Capacity Potential Process Heat (GWh yr^{-1})	180.7	-
Generation Commissioning (GWh yr^{-1})	6.2	-
N Demand (MVA)	340	-
N Generation (MVA)	300	-

Table 6.10 (continued): Example POC input specifications, including technical and demand-related parameters for selected nodes.

Specifications	Example 1	Example 2
Wholesale Electricity Price 2022 (NZD MWh^{-1})	118.13	-

6.8 Green Methanol as a Sensitivity Case

As of 2025, Methanex remains New Zealand's sole methanol producer. Methanex is operating two plants located at Motonui in the Taranaki region, which is the country's main centre for petrochemical and energy-intensive industries. Methanol production at Methanex which uses the conventional steam methane reforming process, relies on natural gas both as a feedstock and as a source of process heat. Approximately 95% of methanol produced by Methanex are exported to the Asia Pacific region, while only about 5% of manufactured methanol is consumed locally. As the local demand is relatively small, the scale of a green methanol industry is viable to decarbonise methanol production to meet New Zealand's needs through renewable pathways.

Methanex has an ongoing long-term contract with the New Zealand gas market where the latter secures 45% of the country's extracted offshore gas production for methanol synthesis (Methanex New Zealand, 2019). A third of the gas supplied is consumed as process heat, with two-thirds of natural gas being feedstock for methanol production; however, as existing natural gas reserves are nearing depletion, the grey methanol production will no longer be sustainable, even to meet local demands. The current price in 2024 for methanol from Methanex is 600 NZD t^{-1}.

There are 71 potential electricity generation sites which are identified to be candidates for green methanol production. These utility-scale renewable generation sites remain uncommissioned, as most operational sites connect to the grid. Utilising these non-commissioned sites allows for the neglect of grid connection costs in the analysis. Figure 6.22 shows the non-commissioned sites by generation type which are identified for electricity generation.

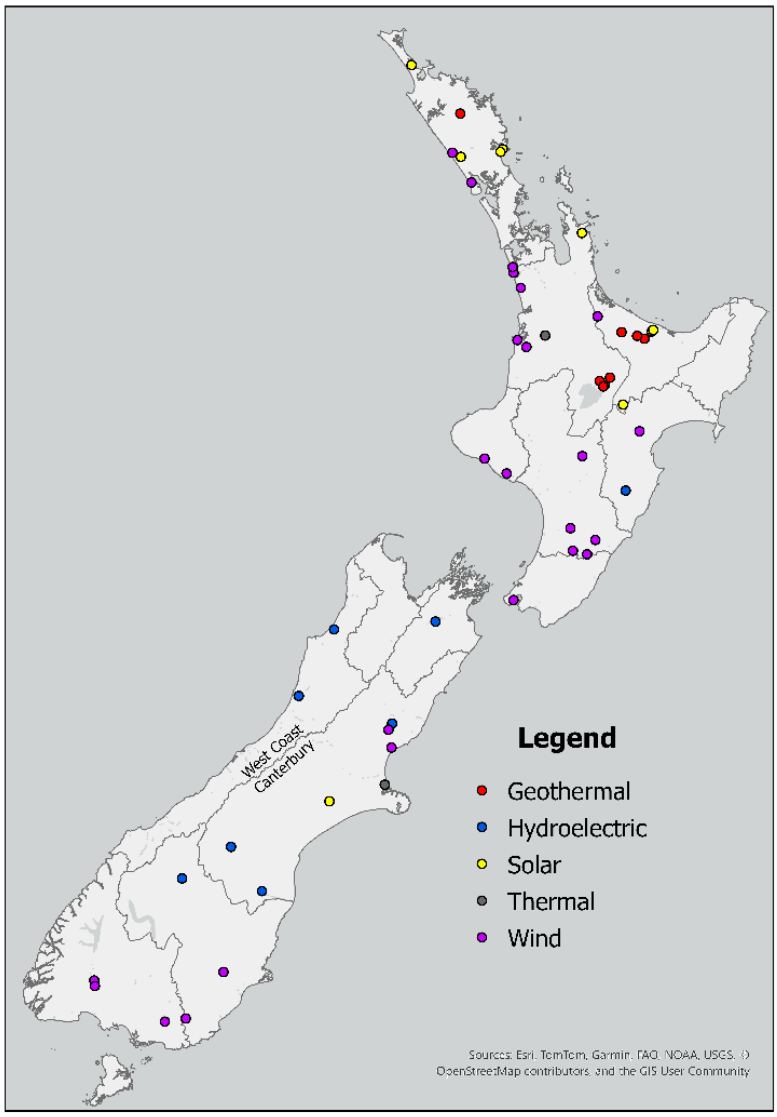


Figure 6.22: Candidate locations for green methanol production, classified by proposed generation type.

6.9 Methods of Geocoding and Distance Calculation

6.9.1 Introduction

This section describes the procedures used to harmonise spatial attributes across all system layers. Because the model represents spatial interactions explicitly, all assets included in scope (industrial sites, electricity generation sites, POCs, and biomass locations) must be assigned with consistent coordinates, geographic coordinate systems, and distance metrics. These spatial attributes determine the feasibility of connections, the distance travelled, and the network representation, which are then used as input for the cost function in the energy model.

The NZTM format developed by Land Information New Zealand (LINZ) is the projection used for official topographic maps for the main islands of New Zealand (LINZ, 2026b). The coordinates format from the Google API uses a global datum which is generally useful as an approach to

analysing and storing coordinates around the world. On the other hand, NZTM format fixes the distortion when coordinates from the Earth's curved surface is being translated into a flat plane. The projected coordinate system uses the format of Easting (E) and Northing (N) with units in meters instead. Thus, it is shaped like a cylinder which depicts elongated areas to account for the difference in latitude, suited for the extent of the New Zealand region. Moreover, longitude distortion is also fixed in the projected coordinate system, which helps maintain accurate spatial relationships and measurements across the New Zealand region. The location based on meters instead of a datum format-based coordinate is suited for GIS, surveying and engineering studies, particularly for energy modelling.

As a summary, the preprocessing tool developed involves the use of the Google Maps Platform to return coordinates, followed by ArcPy and ArcGIS Pro for spatial analysis and data generation into the energy model.

The main spatial layers used in the New Zealand are summarised as:

- **Factory sites (CSV)**: compiled version for Process Heat Database, harmonised from Heat Plant Database and RETA database.
- **Electricity generation sites (CSV)**: compiled and pre-processed from the Electricity Authority database, generation company information, and Open Infrastructure Map.
- **Biomass resource layer (SHP)**: represented as polygons describing land use and forest areas, later converted to point features for distance calculations (Ministry for the Environment NZ, 2024).
- **Road network layer (SHP)**: used to constrain distance calculations to realistic travel routes rather than straight-line distance for biomass travel (LINZ, 2026a).
- **Point of connections (SHP)**: used for spatial allocation of electricity demand and generation and for estimating line distances between connected nodes (Transpower New Zealand, 2025b).
- **Transmission assets (SHP):** used for estimating distance between two or more already established POC connections within the National Grid (Transpower New Zealand, 2025b).

Application of the Data Accumulation Tool on Biomass-Factory and Biomass Generation Connections

After obtaining the geospatial attributes for individual sites, distance calculation is performed using the proposed approach between industrial sites and generation sites with biomass centres. The distance calculation again can be made by following the general methodology which calculates the distance between each entry in both industrial sites and generation sites to biomass databases. However, key work involves creating a polygon layer that represents the road networks, which serves as a "barrier" to limit the search space for distance measurements while ensuring compliance with the valid geographical boundaries of New Zealand.

In Figure 6.23, the source nodes represent the biomass reference point, and the raster surface represents travel distance from that source within the road-constrained barrier. Distance values are then extracted for destination nodes and recorded in meters. All retuned distance values will

be saved in a spreadsheet that records the connection links between biomass-factory and biomass-generation interactions. The example output format for the biomass-factory sites and biomass-generation sites will be displayed in Table 6.11 and Table 6.12 respectively.

From the preprocessing outcome, the analysis of factory sites revealed 27,563 potential connections to biomass locations, representing opportunities to supply biomass fuel for industrial process heat demand. As illustrated in Figure 6.24, there are 417 factory connections located within a 25 km proximity to biomass sources, which highlights regions with high spatial accessibility to biomass resources. Additionally, the total biomass transport cost increases proportionally with distance, demonstrating the sensitivity of biomass delivery costs to spatial separation between source and demand centres. Table 6.12 shows the data format from the data accumulation tool implemented using ArcPy.

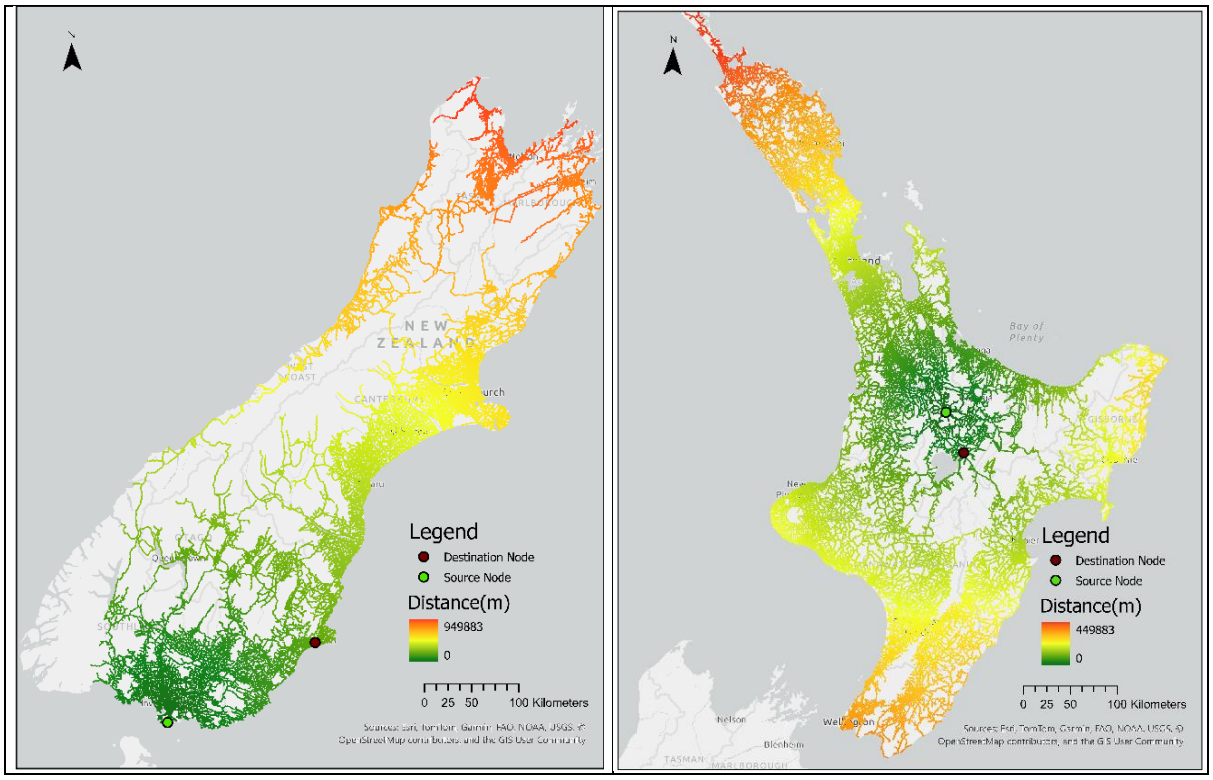


Figure 6.23: Raster output from the Distance Accumulation tool in ArcPy.

Table 6.11: Biomass-generation sites from the iterative Data Accumulation function.

Specifications	Example 1	Example 2
Generation ID	91	303
Generation NZTM-x	1883744	1934570
Generation NZTM-y	5773186	5787726
Biomass ID	132	195
Biomass NZTM-x	1998839	1839014
Biomass NZTM-y	5787367	5672162
Distance	124146	166235
Geographical Boundary	North	North

On the other hand, a total of 16,661 potential connections were identified between generation sites and biomass locations. Based on the frequency distribution shown in Figure 6.25, there are 439 connections within 25 km, 4,010 connections between 25 km and 100 km, and 12,194 connections exceeding 100 km. The output from the pre-processing data indicates that biomass presents a viable opportunity for green methanol production, as it can serve as the carbon source, while hydrogen is supplied through electrolysis powered by nearby electricity-generation plants. Furthermore, biomass can also be used as a substitute fuel in thermal power plants, replacing coal and other non-renewable energy sources for electricity generation.

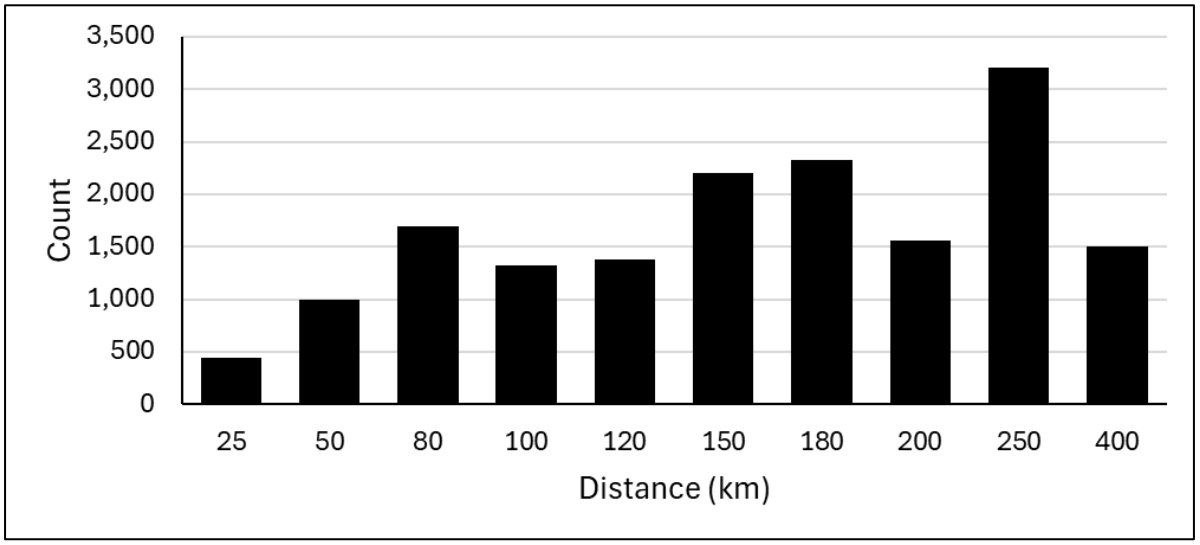


Figure 6.24: Counts of biomass-to-generation connections available as model inputs generated from the Distance Accumulation tool.

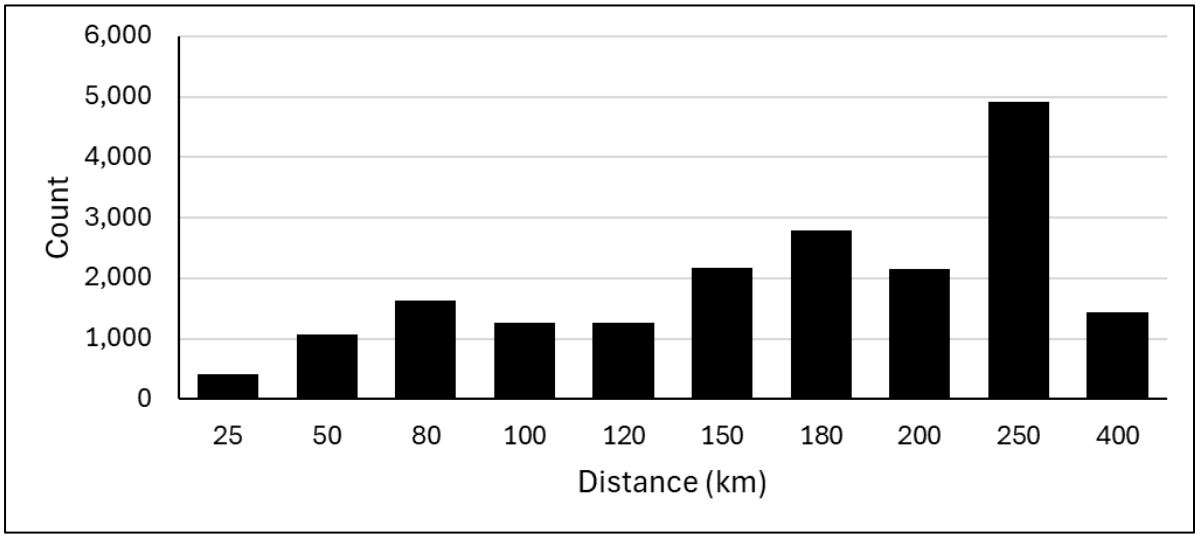


Figure 6.25: Counts of biomass-to-factory connections available as model inputs generated from the Distance Accumulation tool.

Table 6.12: Biomass-factory sites from iterative Data Accumulation function.

Specifications	Example 1	Example 2
Factory ID	1	294
Source NZTM-x	5683076	5607053
Source NZTM-y	1711094	1931381
Biomass ID	78	99
Destination NZTM-x	5485692	5635666
Destination NZTM-y	1859345	1786050
Distance (km)	261.8	174.4
Geographical Boundary	North	North

6.9.2 Distance Estimation between POCs within the Transmission Grid System

Transmission lines can be categorised according to their technical specifications, including voltage level, circuit configuration (single or double circuit), installation type (overhead or underground), and grid capacity rating. These attributes define the transmission limitations represented within the model because distance is required to estimate cost-associated upgrades in the optimisation.

In the New Zealand case study, a challenge arises as each transmission line entry is defined by the line specifications. Databases of transmission line specifications typically define each line entry by the line endpoints (the first and last POC), which can omit intermediate POCs that lie along the physical route. In the New Zealand case study, this required manual linking of adjacent POCs in the geospatial layer to reconstruct a more complete network representation. This process returned approximately 500 bi-directional connections across the transmission grid.

Distance estimation within POCs

Transmission lines can typically be categorised according to their technical specifications. These include parameters such as voltage level, circuit configuration (single or double circuit), installation type (overhead or underground), and grid capacity rating. These assets determine the transmission limitations represented within the model.

The regional transmission authority typically maintains databases containing information on transmission line specifications and the associated network connection nodes. Further pre-processing was required to calculate the distances between POCs. Unless the distance is provided between POC locations is provided, a similar distance estimation methodology is applied across all POC connections, which is the ArcPy-based Distance Accumulation tool. Both the biomass connection distances and POC connection distances were computed using this approach.

However, the methodological implementation differs between both distance calculations. Unlike biomass connections, destinations are set dynamically, where road networks successfully link with generation and factory sites while adhering to the geographical constraints of the routing path. On the other hand, POC connections calculate distances within their own feature layer, where the connections have already been pre-defined. As a result, the POC distance analysis

offers less spatial flexibility compared to biomass routing, which traverses across all possible connection paths.

If the line distances between connected POCs are not given, the distances between two POCs are estimated using an ArcPy-based Distance Accumulation workflow, consistent with the approach used for biomass routing but applied within the transmission grid feature layer. Unlike biomass distance calculations, where destinations are dynamically connected via road networks, POC distance estimation is performed between already defined network connections. A spatial barrier was constructed to remove overlapping areas within the transmission grid feature layer and to confine the raster extent within valid boundaries. POC locations were inserted as point features using their x–y coordinates, and the Distance Accumulation tool was applied iteratively across the set of connected POC pairs. Distances were then obtained by overlaying destination POC points onto the generated raster surface, and the resulting values were recorded as the transmission distance between POC connections. An example of the resulting outputs is shown in Table 6.13.

Table 6.13: Distance estimation within POCs and the Distribution Level.

Specifications	**Example 1**	**Example 2**
Point 1	RTO	ROS
Point 2	ONG	PEN
Grid Name	ARI-ONG-B	PEN-ROS-A
Distance (km)	50.0	9.0
Geographical Boundary	North	North
Capacity Rating (MW)	57.14	411.46
Voltage (kV)	110	110
Line Type	Single	Double

6.10 Verification for Model Inputs

6.10.1 Intermittent Generation

Because the model uses aggregated monthly profiles and 4-hour time blocks, short-term variability is smoothed and differences between simulated and measured series are reduced. Using Taupo WXT AWS wind speed observations from (MetService, 2025), the hourly series comparison over 2009-12-18 to 2016-07-07 (n = 2383) gives an RMSE of 0.672 m s^{-1}. When both series are aggregated to monthly means, the RMSE reduces to 0.002 m s^{-1}. These results support the use of Renewables.ninja-derived wind inputs for the aggregated temporal resolution used in the model (Figure 6.26). The daily (hourly) comparison shows that simulated wind speeds exhibit lower peak magnitudes and less short-term variability than the measured series, reflecting the smoothing inherent in reanalysis-based products.

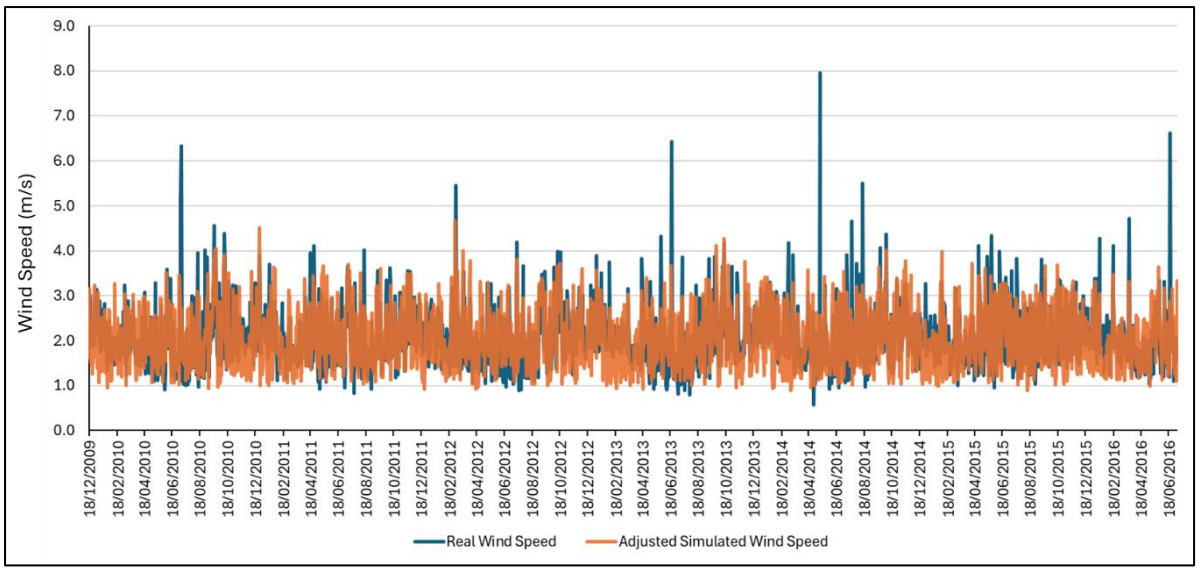


Figure 6.26: Comparison of actual daily wind data from MetService (2025) and simulated wind data from Renewables.ninja from 2009 to 2016 overlaid in the same graphical plot.

Based on the monthly figure, the seasonality effect on the wind speed by months can be captured in Figure 6.27. This indicates that the simulated aggregated monthly wind speed from MERRA-2 will be used for the energy model as it reflects the similar trend of the measured value.

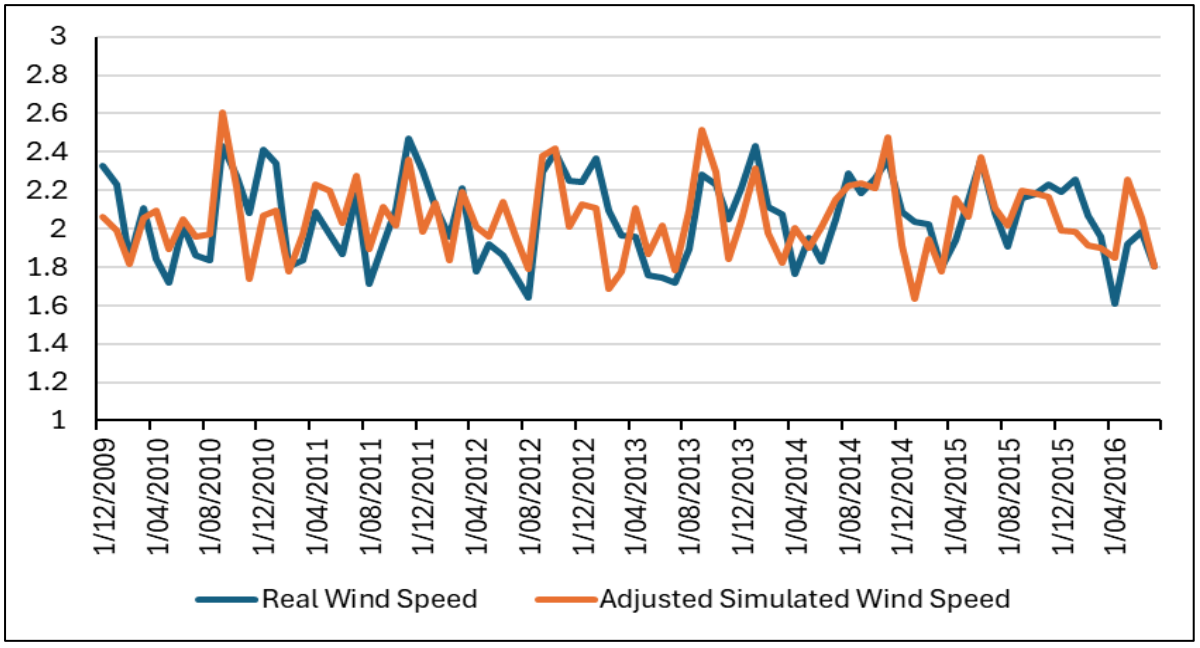


Figure 6.27: Actual aggregated monthly wind data from MetService (2025) and simulated wind data from Renewables.ninja from 2009 to 2016 overlaid in the same graphical plot.

A verification check was conducted for the solar generation data. The simulation data obtained from MERRA-2 Global were extracted for a five-year period. The data was then compared against the actual recorded values from Solarview (NIWA, 2025). The real-world data used for validation were collected from the Lodestone Energy Edgecumbe Solar Farm, representing measured average diffuse irradiance, as shown in Figure 6.28. As illustrated in Figure 6.28, the simulated and measured data exhibit similar solar patterns, indicating that the synthetic dataset is reliable

and can be used in the model without compromising result quality. The agreement between datasets was evaluated using the root mean square error (RMSE) over the validation period, yielding a value of 15.6 kWh m^{-2}. In addition, Figure 6.29 presents the actual average solar irradiance data: (a) seasonal variations shown in a line plot, and (b) hourly variations across months shown in a heatmap. These visualisations highlight how seasonal and monthly differences influence irradiance levels throughout the day.

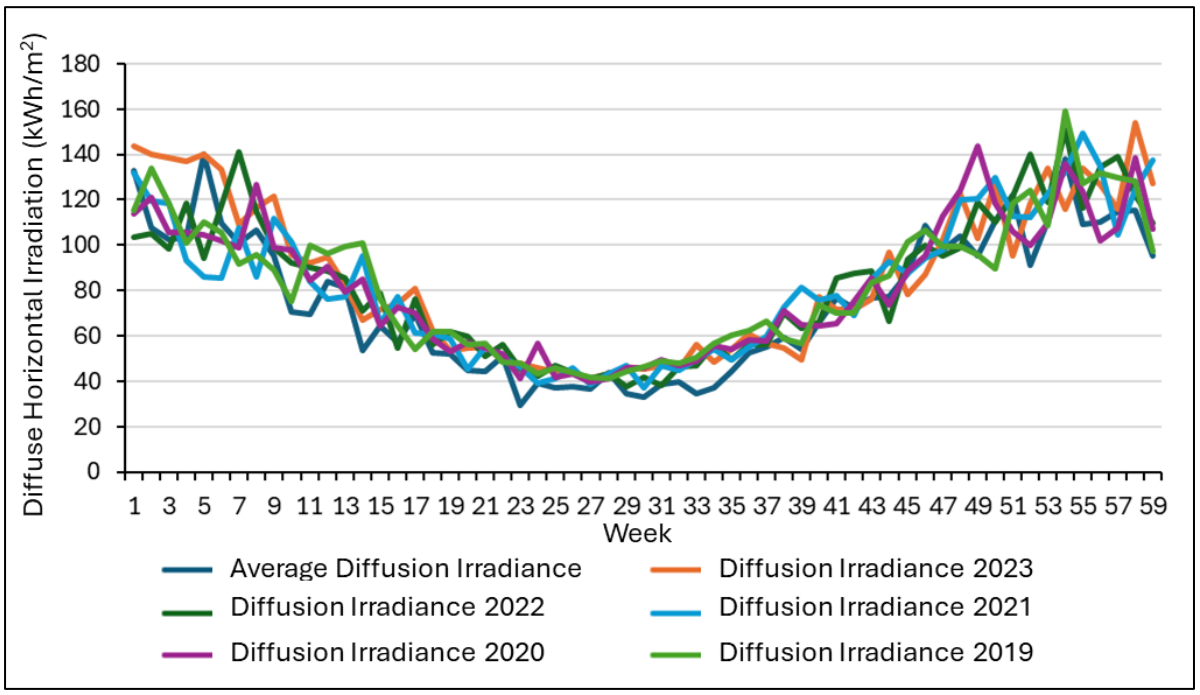
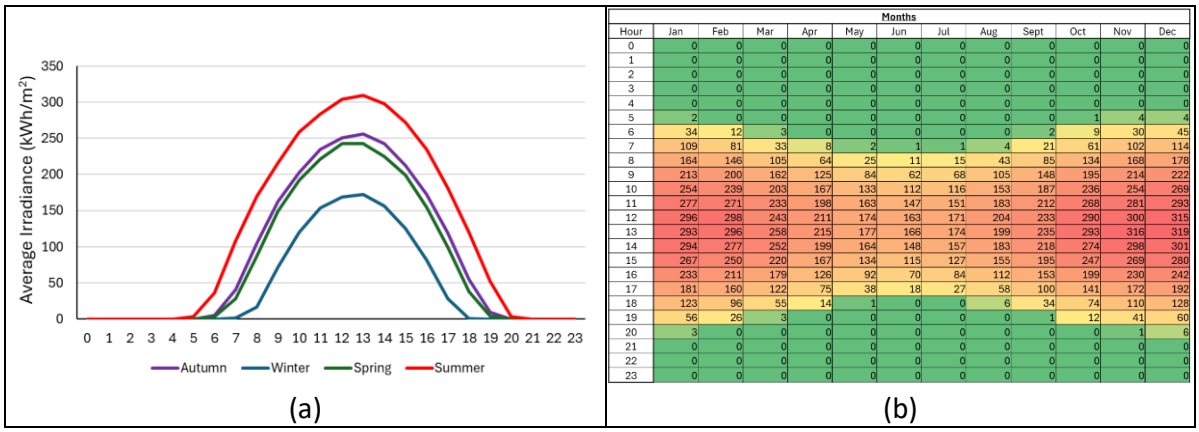


Figure 6.28: Actual average solar irradiance data from Solarview (NIWA, 2025) and simulated data from Renewables.ninja in weeks overlaying in the same graphical plot.



Months

Hour	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0
5	2	0	0	0	0	0	0	0	0	1	4	4
6	34	12	3	0	0	0	0	0	2	9	30	45
7	109	81	33	8	2	1	1	4	21	61	102	114
8	164	146	105	64	25	11	15	43	85	134	168	178
9	213	200	162	125	84	62	68	105	148	195	214	222
10	254	239	203	167	133	112	116	153	187	236	254	269
11	277	271	233	198	163	147	151	183	212	268	281	293
12	296	298	243	211	174	163	171	204	233	290	300	315
13	293	296	258	215	177	166	174	199	235	293	316	319
14	294	277	252	199	164	148	157	183	218	274	298	301
15	267	250	220	167	134	115	127	155	195	247	269	280
16	233	211	179	126	92	70	84	112	153	199	230	242
17	181	160	122	75	38	18	27	58	100	141	172	192
18	123	96	55	14	1	0	0	6	34	74	110	128
19	56	26	3	0	0	0	0	0	1	12	41	60
20	3	0	0	0	0	0	0	0	0	0	1	6
21	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0

Figure 6.29: Actual average solar irradiance data: (a) seasonal variation shown in a line plot, and (b) hourly distribution by month presented as a heatmap.

6.10.2 Baseline energy demand

The MBIE's Energy Report (MBIE, 2024b) clearly outlines the total energy balance for each sector. It constitutes the total energy demand for each industrial sector, such as food processing, textiles, wood pulp and paper, chemicals, non-metal, metals and others. The fuel type is also clearly outlined in the database, which includes renewable fuels, non-renewable fuels and electricity.

Sector-level electricity shares from MBIE are used as consistency checks and to inform site-level electricity allocation where site-metered values are unavailable. Therefore, based on the intensity of the energy used per factory, the electricity use on a level which can be approximated based on the compiled Process Heat Database from various sources.

Table 6.14 shows that sectors such as basic metals and textiles have a high share of electricity use. With 88% for basic metals and 59% for textiles-based industries. Sectors such as food processing, non-metals, chemicals, wood, pulp and paper have higher energy-based fuel use, as the processes mainly comprised the heating and cooling of the throughput materials, with food processing using 74% of non-renewable fuels, 84% from non-metals and 90% from chemicals, while the share of energy consumption in the wood, pulp and paper industries is 1% coal, 10% natural gas and 74% solid biofuels with 16% electricity.

Table 6.14: MBIE Energy balance table based on ratio (MBIE, 2024b).

Industry	Coal	Oil	Natural Gas	Renewables	Electricity
Mining	0.00	0.76	0.01	0.00	0.23
Food Processing	0.27	0.00	0.47	0.00	0.26
Textiles	0.02	0.00	0.40	0.00	0.59
Wood, Pulp, Paper and Printing	0.01	0.00	0.10	0.74	0.16
Chemicals	0.00	0.00	0.90	0.00	0.10
Non-metallic Minerals	0.50	0.00	0.34	0.00	0.16
Basic Metals	0.01	0.00	0.11	0.00	0.88
Mechanical/Electrical Equipment	0.00	0.00	0.34	0.00	0.66
Building and Construction	0.00	0.79	0.05	0.00	0.16
Unallocated	0.00	0.43	0.01	0.22	0.34

6.10.3 Baseline electricity demand by POC

Estimated electricity POC demand is constructed by summing allocated residential, commercial, industrial, and BEV components and compared against measured POC demand for 2022. To validate the estimated assumptions and actual electricity results, four key POCs which exhibit different load characteristics are selected for validation. The assumptions include the increase in electricity demand corresponding to population growth, as described in Section 4.4.5, and the monthly electricity demand for industries based on the industrial load curve developed in Section 4.4.2. The actual electricity price are obtained from Electricity Authority (Electricity Authority, 2025).

The POCs are Paraparaumu (PRM) in Wellington, Edendale (EDN) in the Southland region, Glenbrook (GLN) in Auckland and Penrose (PEN) in Auckland. POC situated in Paraparaumu is predominantly residential and commercial, with little to no industrial activity. Edendale is a town mainly focused on dairy processing. Glenbrook has a steelmaking industry, and Penrose in

Auckland is a mix of industrial, commercial and residential sectors. Figure 6.30 shows the makeup of the assumed electricity demand by category.

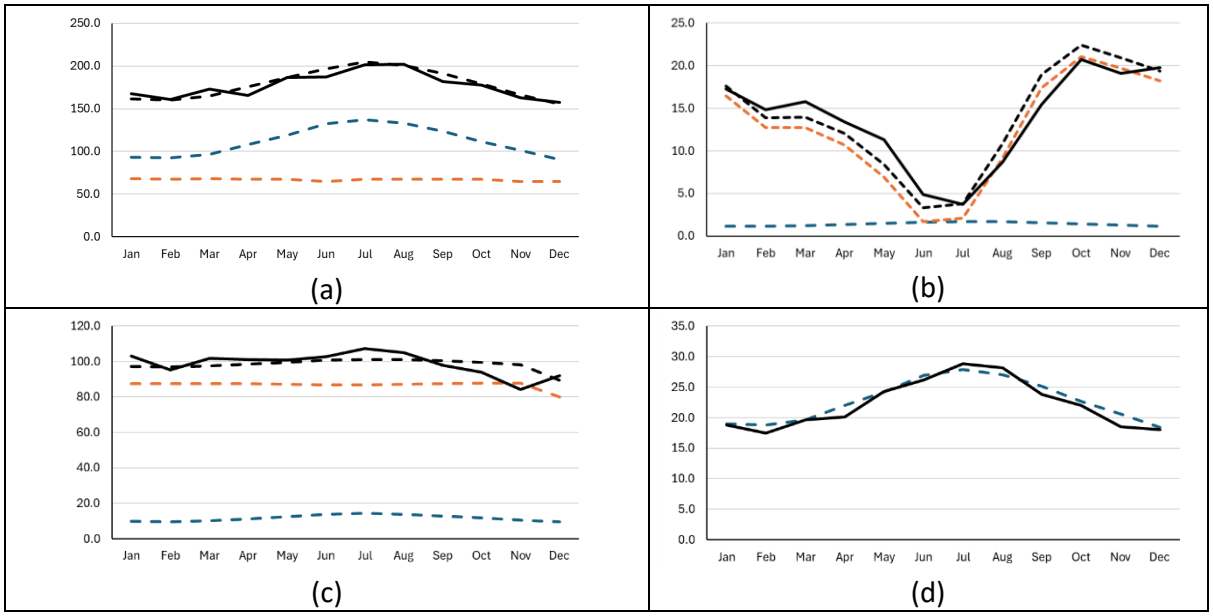


Figure 6.30: Monthly electricity demand profiles in GWh for (a) Penrose, (b) Edendale, (c) Glenbrook, and (d) Paraparaumu, showing total actual electricity demand (black solid), estimated total electricity demand (black dotted), estimated industrial demand (orange dotted), and residential and commercial demand (blue dotted).

In all four examples, it can be observed that the estimated demand could approximate the total electricity demand with a low error. However, the randomness and erratic changes to the months in each POC are not captured. For the Paraparaumu POC, the estimated trend is fairly accurate with the curve, where the electricity peaks from June till August. The trend validates that the residential trend and population per ICP estimates correspond well to the actual residential demand. The estimated annual demand is 272.5 GWh, whereas the actual demand was 266 GWh, with a difference of 6.5 GWh. As for the Edendale POC, the milk curve pattern can be seen existing on the trend. This is resulted from the existing Fonterra's Edendale site, which is one of the largest dairy factories in New Zealand, which processes up to 15 million litres of milk daily (Fonterra NZ, 2022). The electricity is topped-up by an average of 1.4 GWh per month by non-industrial consumers, including 640 residents in the region (Stats NZ, 2024). In summary, the estimated total electricity consumption is 165.4 GWh and the actual is 164.8 GWh, with a marginal error of 0.8%.

As steelmaking dominated Glenbrook POC, the estimated total electricity from industrial use was 1040.6 GWh, supported by consumer use from non-industrial of 139.7 GWh. Comparing with the recorded electricity use, the actual demand is 1184.2 GWh, which has a difference of 3.9 GWh in estimates. The industrial curve expects the metal industry to have consistent operational hours. However, the drop in consumption might be due to scheduled plant shutdown or maintenance which is timed differently across all metal sites.

The next comparison is made for a POC which have large industrial and non-industrial users. The Penrose POC is typically a large energy sink in the Auckland region with a recorded electricity use of 2125.4 GWh in the year 2022. Based on the estimates, the non-industrial use accounts for 58%

of the total electricity share during summer and up to 67% during winter months. By accounting for the non-industrial and industrial use separately before adding them together, both industrial-based trends, driven by the sector, and non-industrial trends can be recorded for the multiperiod patterns.

The industrial electricity use in Penrose is approximated to be consistent, as it is typically made up of food (4), metal (3), non-metal (3), chemical (3) and others (3). Figure 6.31 shows the makeup of the industrial sectors and the percentage of electricity share for each sector. The electricity is consumed the most in the non-metal and chemical plants, which are processing glass, plastics, polystyrene, construction, and waste management. As most industrial activities connected to Penrose are not seasonally affected, the electricity demand remains consistent, whereas the residential side electricity use follows a different monthly pattern.

In summary, the projected electricity is 46.34 TWh yr^{-1} based on the projection assumptions. Additionally, end-use demand of 25.9 TWh yr^{-1} is required if a scenario of 100% electrification of industrial process heat is committed.

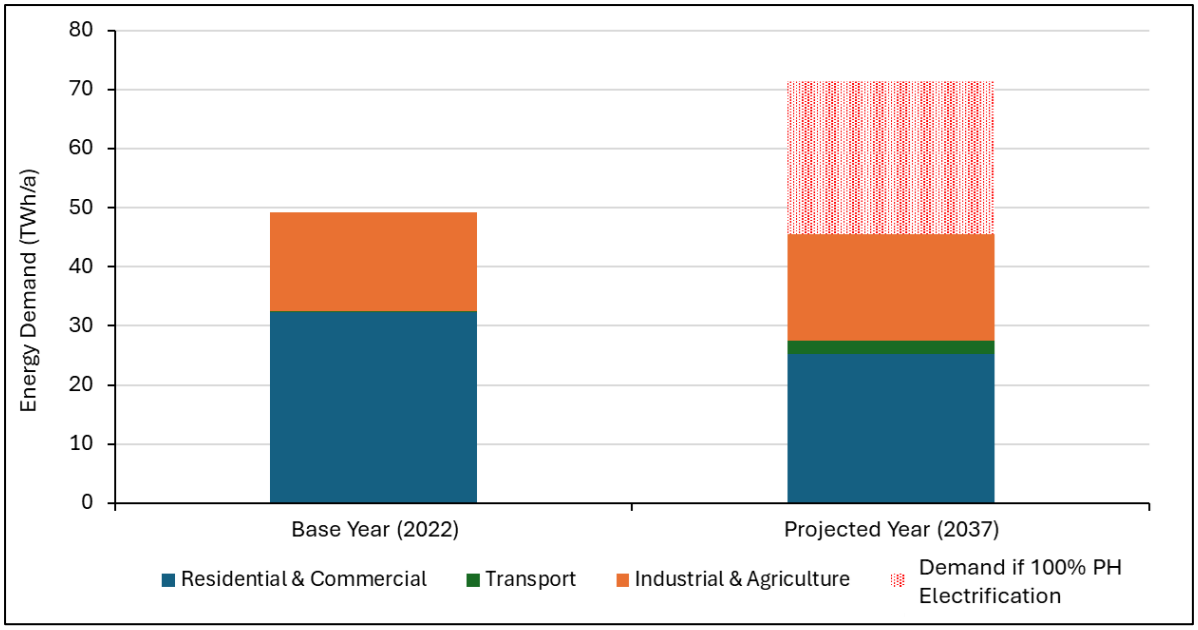


Figure 6.31: Comparison of reference year vs projected year for process heat optimisation.

6.10.4 Projection scaling assumptions (population and BEV uptake)

As the quantity of residential and commercial demand is affected by the number of populations connected to a POC, the population projection is used to assume the electricity use from residential demand. As of September 2022, there are approximately 5.1 million residents in the country. The statistics from Stats NZ show that there will be an annual average of 1.0% population growth up till 2037, when the population will be bumped up to 6.0 million inhabitants. The change of population from the projected year is estimated by Stats NZ to be 14%. Therefore, to account for the population rise in the country leading to more electricity generation being required, the

electricity consumption for the projected year from residential and commercial has to be increased in proportion with growth. At this instance, it is assumed that the population growth is evenly distributed across POC, as it poses a challenge to account for internal migration and unequal population booms in certain regions catalysed by governmental policies and job opportunities. The national electricity consumption increases by 3.1 TWh, up to 25.2 TWh.

Referring to the increase of market share for electric vehicles, the number of electricity required for charging electric vehicles will be set at 2.3 TWh yr^{-1}, which is a sharp rise from the initial 0.23 TWh yr^{-1}. Figure 6.32 illustrates the electricity increase by regions across New Zealand by increments. Regions such as Auckland, Wellington and Canterbury witness a large increase from the base year. This increase resulted from the high population share. On the other hand, industries such as the Southland region are mainly industrial demands, driven by a strong presence of metal, dairy, meat and wood processing sectors in the region. These inputs set the parameters for electricity demand growth, which does not directly correlate with decarbonisable industrial process heat. However, they remain important for modelling the projected year, particularly in the grid reinforcement analysis presented in Chapter 8.

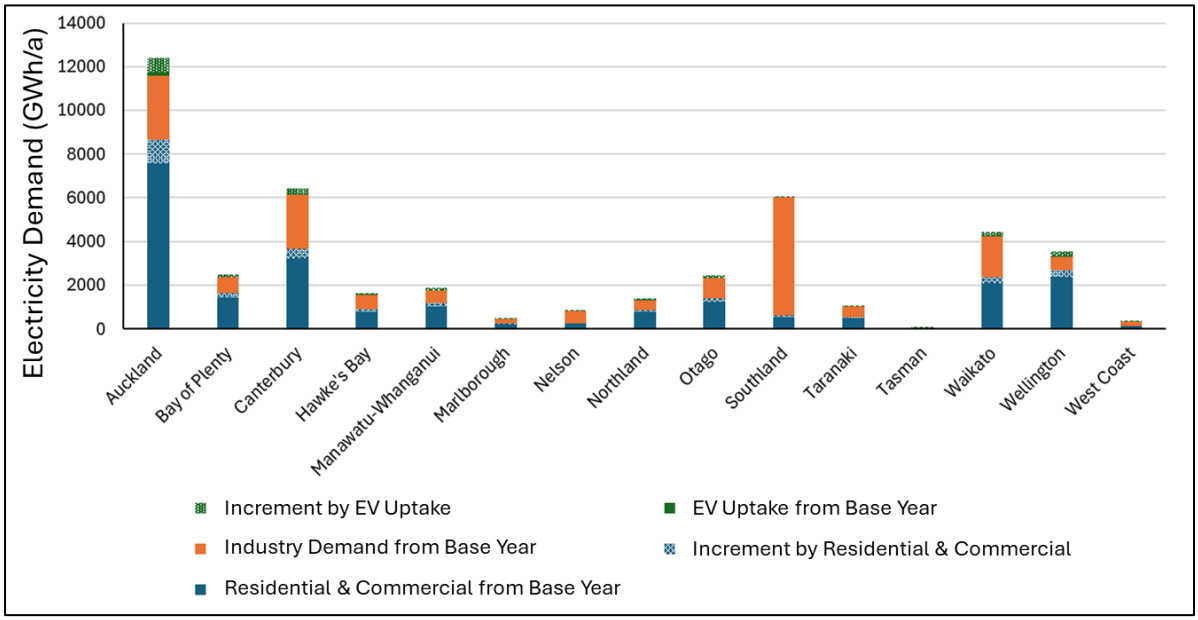


Figure 6.32: Projected sectoral electricity demand by region in 2037, excluding electrification of decarbonisable industrial process heat.

6.11 Chapter Summary

In this chapter, a comprehensive set of data inputs and modelling dimensions was defined. This includes industrial process heat demand, electricity generation assets, biomass resource availability, and electricity grid infrastructure. These elements are core to representing the interactions between supply, demand, and interconnections within the modelling framework. A contribution to the New Zealand industrial process heat case study is made by identifying and compiling all decarbonisable major process heating sites with fuel-switching demand, operational industrial curves based on processes, current boiler type, current major fuel type, connected POC, coordinates, demand by temperature bands, and other characteristics. A particular emphasis was placed on ensuring consistency and reconciliation of datasets, given the harmonisation of multiple sources such as EECA, MBIE, Transpower, the Electricity Authority, and other national databases and reports.

The chapter also introduced a site-level representation of industrial process heat demand, including temperature classifications and sectoral distributions. This enables the model to capture technology compatibility and fuel-switching pathways with greater resolution compared to aggregated approaches. In parallel, the representation of electricity generation and grid infrastructure, including Points of Connection (POCs), transmission, and distribution layers, allows the model to evaluate the impact of electrification on network capacity and expansion requirements.

A key contribution of this chapter is the integration of geospatial modelling techniques, including geocoding and distance calculations. The intention is to establish realistic connections between biomass resources, industrial sites, and electricity infrastructure. These spatial relationships are critical for capturing logistical costs, resource availability, and infrastructural constraints, which influence the outcome of the system configuration.

Finally, the chapter presented verification and validation of model inputs, ensuring that the compiled datasets and assumptions are consistent with national energy statistics and realistic system behaviour. This step is essential to ensure that the input parameters are accurate, thereby improving the credibility and reliability of the modelling results demonstrated in the subsequent chapters. Overall, this chapter operationalises the modelling framework by applying a New Zealand case study, bridging the gap between theoretical mathematical model development and practical application.

The subsequent three chapters apply this case study across all three simulation modes, deterministic multiperiod, sequential and Monte Carlo, which altogether represent a cumulative argument instead of three separate analyses. Chapter 7 first establishes the least-cost fuel-switching pathway under a set of deterministic assumptions, which isolates the influence of temporal variation in demand and resource availability on technology selection. Chapter 8 relaxes the assumption that electricity supply is exogenous by introducing a sequential formulation that allows electrification decisions and grid reinforcement costs to interact. Chapter 9 relaxes the assumption that input parameters are known with certainty, by applying Monte Carlo sampling to assess whether the pathways identified under Chapters 7 and 8 remain robust, or whether they

are artefacts of a specific set of deterministic assumptions. No single mode alone can establish that a fuel-switching pathway is simultaneously cost-optimal, consistent with its own electricity system requirements, and robust to parameter uncertainty. Only by progressing through all three does the thesis establish this jointly.

Chapter 7: Multiperiod Optimisation

7.1 Introduction

This chapter presents the deterministic multiperiod optimisation framework used to evaluate fuel-switching pathways for industrial process heat under time-varying conditions. Building on the case-study inputs developed in Chapter 6, it applies the energy modelling framework to identify least-cost technology configurations, operating levels, and capacity requirements across the full modelling horizon. The multiperiod structure allows the model to capture how seasonal demand patterns and temporal variation in electricity supply conditions influence the feasibility and cost of electrification and biomass-based heat provision. The chapter focuses on the factory-side optimisation problem for 2037 demand levels, using 2022 as the reference year for prices and technology cost inputs, with all values expressed in real 2022 NZD. The chapter focuses on the least-cost fuel-switching pathway under deterministic conditions, which forms a baseline result where the sequential and Monte Carlo analyses in Chapter 8 and 9 are tested.

The scope of the chapter is focused on the costs incurred by the factory side for process heat decarbonisation. This analysis includes the capital and operating costs of technology switching, along with the cost of purchased electricity and fuels required to meet the required heat demand for each time step. In this case study example, electricity costs are applied as an energy volume based on sales cost, which reflects business tariffs dominated by a FPVV structure derived from wholesale price patterns. The electricity cost calculation is intentionally simplified and does not include investment-based representations of new generation build or grid upgrade costs.

Electricity generation is included only to preserve temporal consistency and system feasibility in the background, ensuring that baseline electricity demand from residential, commercial, and industrial consumers can be met across time periods, but it is not treated as an endogenous build decision or a construction cost component within the objective. The exception is the green methanol proxy, where electricity generation costs are explicitly incurred as part of the production pathway.

This separation of cost components is intentional. Chapter 7 provides a clear model overview of how a single deterministic multiperiod optimisation translates assumptions about demand patterns and technology options into time-dependent fuel-switching decisions. Chapter 8 then builds on the same foundation by introducing grid asset limitations and upgrade requirements and by using a sequential deterministic formulation to quantify the additional system costs associated with network constraints and reinforcement.

7.2 Deterministic Multiperiod Solution Approach

In a multiperiod approach, each time step is represented by discrete monthly periods, and the time index is applied uniformly across the system. Time step is represented by a finite set of periods $t \in T$, where $T = \{t_1, t_2 \dots t_{12}\}$. The time is represented by average operating conditions for each month.

The projection year of 2037 includes scaling of electricity demand. However, costings such as electricity prices, technology capital costs and fuel costs are fixed at real 2022 values, as the focus is on structural and operational feasibility, not price forecasting. Moreover, the technology selections used are mature technologies. However, sensitivity analysis will be discussed in the later section.

The model, including for future chapters, identifies static and time-varying definitions deliberately. Time-invariant decisions such as technology selection, which is a disjunctive choice for structural optimisation, and installed capacity where accounts for the least cost for the overall energy system. On the other hand, time-varying features include heat demand, electricity consumption, fuel consumption, and the operational level of technology for each time step. Therefore, it is important to note that the technology capacity is installed once and shared across all time periods, meaning there will be no resizing with each time step change. However, the only change is the inflow and outflow of energy within the operating capacity, which is varied for every time step.

As part of representing seasonal variation in monthly demand profiles, several key parameters are addressed, including intermittent electricity generation, electricity prices, POC demand, and allowable industrial process heat for decarbonisation. Intermittent electricity generation data obtained from Renewables.ninja are aggregated into average monthly capacity factors to be used in the model. The allowable process heat demand for decarbonisation at industrial sites is generalised using sector-specific monthly industrial demand profiles, which are derived from multiple data sources and supported by the justifications made in earlier sections.

In this energy model, perfect foresight modelling is employed. Essentially, the parameters are defined as model inputs with zero uncertainty. Hence, by knowing the parameters for the next time steps, the problem is solved simultaneously, the assets in the model are sized flawlessly, and the system cost is minimised optimally (Maier *et al.*, 2025). Perfect foresight is applied in the energy model due to its size and complexity, where a deterministic solution is used to obtain a global optimum solution structure (Payet-Burin *et al.*, 2021). Despite no uncertainty represented in this chapter, sensitivity analysis will be described in the later section, and elements of stochastic Monte Carlo optimisation, which deals with uncertainty, will be applied in the energy modelling framework.

Although the optimisation model is implemented using the P-graph framework in Python, the formulation is presented in mathematical form for clarity and generality. As P-graph is a good representation for a superstructure for an energy model and it is effective to demonstrate the energy flow and capacity selection, it is, however less intuitive when communicating constraint logic and objective relationships in a thesis context. Therefore, by expressing the model in an algebraic format, the mathematical formulation serves as a more transparent and general description.

7.3 Model Formulation

7.3.1 Sets

The multiperiod optimisation for the energy model is structured around a series of sets which represent the explicit indices within the biomass, electricity generation, transmission and factory layers. These are shown below:

Time period T denotes the number of time steps, t in the model. Since the temporal resolution is monthly, the $|T| = 12$.

The electricity generation layer has two sets: electricity generation sites J and generation types E. The set for electricity generation sites J represent the set which constitutes the number of electricity generation nodes j, with $|J| = 291$ sites. A total of eight electricity generation types for the set E, where $E = \{Utility\ Solar,\ Local\ Solar,\ Wind,\ Hydro,\ Thermal,\ Geothermal,\ Cogeneration, Bioenergy\}$.

The biomass layer constitutes two sets: biomass resource sites K and biomass resource type M. The set for Biomass resource sites K describes the number of biomass locations available within the geographical boundary k with $K = \{k_1, k_2, \dots, k_{214}\}$ whereas there are 5 elements in the resource type set, W which are $W = \{Inforest\ harvest\ residue,\ Klogs,\ Sawmill\ Chip,\ Straw\ and\ Stover,\ Pellets\}$.

The set for POC nodes is defined by I, with a total of $|I| = 217$ different locations. As for the connection network within neighbouring POCs, a second set is required. The pairs (i, u) are denoted by set A^{POC}, where $|A^{POC}| = 652$ individual connections.

For the industrial layer, three sets are defined, namely, industrial site location L, heating technology P and temperature profile Q. The number of industrial sites l is 298, where $l \in L$ and there are seven types of heating technologies $P = \{LTHP, MTHP, HTHP, EB, BB, BFB, CB\}$. The temperature profile is described by q, where $q \in Q$.

Figure 7.1 shows the model superstructure consisting of sets and indices to represent each component in the mathematical formulation.

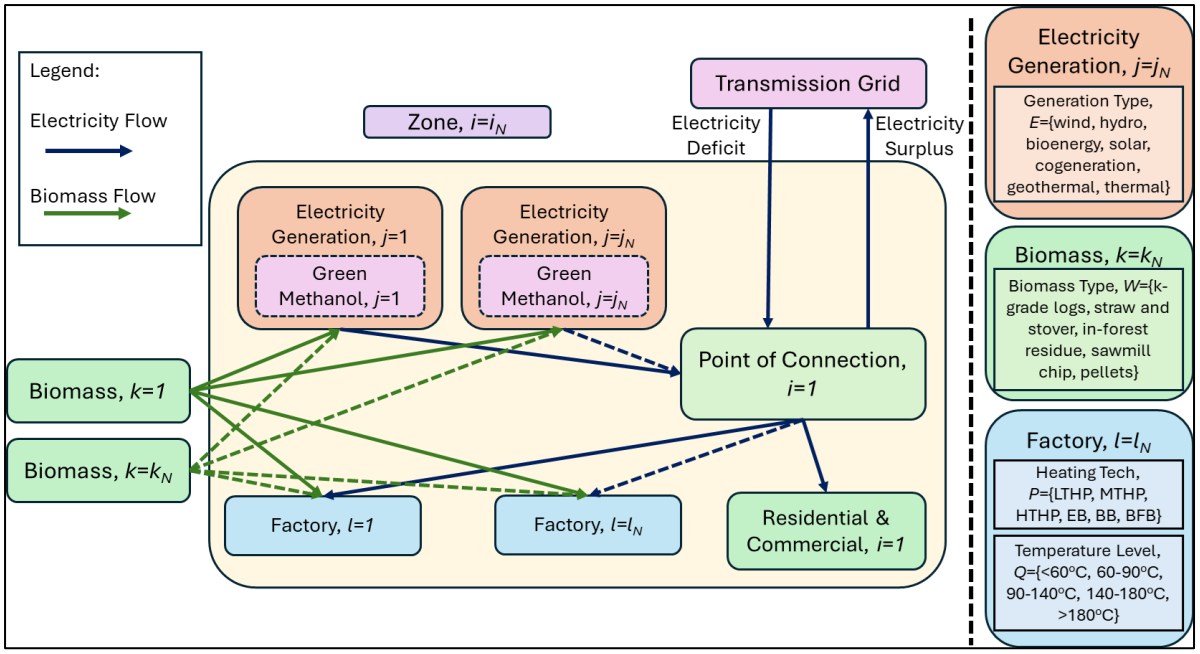


Figure 7.1: Superstructure model of sets and indices in the mathematical formulation.

7.3.2 Parameters

Several key specifications have been identified previously before the model is simulated. These include existing sites or technology capacities, which are pre-processed or predefined for the model. These include parameters which have binary and continuous characteristics. Typically, the costings for sizing and technologies per unit value serve as cost parameters aimed at minimising the overall energy system. For example, the capacity factor and the installed capacity for existing technologies are obtained from external sources. For the POC layer, electricity demand for each POC has been previously assumed and used as an exogenous input. Potential connections between layers such as biomass and factories, biomass to generation, generation to grid and grid to factories are defined by geospatial tools, along with the cost-associated distance being developed.

Biomass availability is set up as a resource constraint and has to be defined so that the utilised biomass amount does not exceed the biomass allowable limit. On the factory side, temperature demand for each site has been studied and compiled, which includes the demand for each temperature level. The coefficients of performance (COP) and efficiency for the technologies are determined in advance. Moreover, allowable capacity for retrofitting coal boilers with biomass resources in each industrial site is defined as a pathway for fuel-switching in industries.

7.3.3 Decision Variables

There are several key decision variables that serve as the primary indicator for all layers. These variables govern technology adoption, capacity sizing, and operational dispatch across time periods. Other additional variables are included for constraint expression, flow balance and

connectivity between variables. These variables and constraints are not discussed in detail in this chapter but included in Appendix C as these do not express key insights from the optimisation output. The main decision variables which are grouped by layers are as follows:

Industrial Layer

Industrial process heat layer represents process heat technology selection, capacity sizing, energy input (fuel) and energy output (heat) for each temperature band. These variables include electricity flow $F_{l,t,p}^{e,in}$ for technology p in site l at time t, biomass flow $F_{l,t,p}^{bio,in}$ for technology p in site l at time t, heat outlet from electric $F_{l,t,p}^{e,in}$ heating technology p for temperature band q in location l at time t and heat outlet from biomass $F_{l,t,p,q}^{b,out}$ heating technology p for temperature band q in location l at time t. This also includes heat supplied $H_{l,t,p,q}^{sup}$ from technology p, for temperature band q in location l at time t.

Biomass Layer

Biomass variables refer to the amount and type of resource used for each biomass node and transport of biomass to various layers. These include biomass $B_{k,w}$ used, biomass transport to methanol plants $B_{k,j,w}^{bio,MeOH}$, biomass transport to industrial locations $B_{k,l,w}^{bio,fac}$ and biomass transport to generation sites, $B_{k,e,j,w}^{bio,gen}$.

POC Layer

The POC variables consist of electricity deliveries within the grid and between layers. These include electricity flow $F_{j,i,t}^{e}$ from generation site j to POCs i at time t, electricity flow $F_{i,u,t}^{POC}$ from POC i to POC u and electricity flow $F_{i,l,t}^{e}$ from POC i to factory location l at time t. Moreover, the electricity flow $F_{j,t}^{e}$ to meet demands which are projected non-industrial process heat in POC i at time t.

Generation Layer

Generation-related variables represent installed capacity, operational dispatch and storage behaviour across time periods. These include planned electricity generation capacity $Cap_{e,j}^{e}$, $\forall(e,j) \in \varepsilon^{planned}$, electricity generated $F_{e,j,t}$ by technology e in location j at time t, as well as storage state variables $S_{e,j,t}^{e,st}$, which determine the available electricity supply delivered $F_{j,t}^{gen}$ to the POC i at time step t.

Methanol Layer

Methanol related variables include the amount of electricity and biomass flow entering the methanol site j at given time t. These include input variables of electricity $F_{j,t}^{e,MeOH}$ and biomass $F_{j,t}^{bio,MeOH}$. Methanol production flow $G_{j,t}^{MeOH}$ in location j at specific time t and sizing capacity Cap_{j}^{MeOH}.

7.3.4 Objective Function

As the optimisation problem is presented in P-graph and the problem is transposed into the Pyomo solving framework, the model inputs and structural relationships are defined in the P-graph framework. Therefore, P-graph does not rely on explicit algebraic sets or indices like Pyomo; hence, in the porting tool, the sets and expressions are demonstrated in Section A are generic, all based on the type of nodes and arcs within the P-graph framework.

However, despite P-graph being an intuitive visual representation framework for representing the system superstructure, it is hard to be expressed in formulations with the absence of sets and indices. Hence, all the expressions and decision variables are represented in a way where the algebraic mathematical formulation format is denoted for the problem.

The purpose of presenting the objective function and constraints in algebraic form is to clarify the cost structure and decision logic embedded in the P-graph model and to provide a bridge between graph-based PNS and conventional algebraic energy system modelling.

In the energy model, the objective function is to minimise the overall system cost (Equation (7.1)). It is also linked with a few supporting equations, such as Equation (7.2) to Equation (7.4) which derive the terms from the energy model, are being formulated as:

$$\min CC_{Resource} + CC_{Op} - CC_{Product} \quad (7.1)$$

$$CC_{Resource} = CC^{bio,res} \quad (7.2)$$

$$CC_{Op} = CC^{MeOH} + CC^{e,st} + CC^{bio,fac,trans} + CC^{bio,MeOH,trans} + CC^{ht} + CC^{salescost,PH} \quad (7.3)$$

$$CC_{Product} = Rev^{MeOH} \quad (7.4)$$

7.3.5 Cost breakdown

In order to support the transparency in the objective function, total system cost is broken down into several major cost components. The following section only describes the cost components which are central for the study, which includes further cost components that can be found in the Appendix D section.

Biomass Transport cost

$$CC^{bio,fac,trans} = \sum_{k\in K}\sum_{w\in W}\sum_{l\in L} C_{k,l}^{trans} T_{k,l,w}^{bio,fac} \quad (7.5)$$

Biomass Resource cost

$$CC^{bio,res} = \sum_{k\in K}\sum_{w\in W} C_{w}^{bio} B_{k,w} \quad (7.6)$$

Process Heat Technology Cost

$$CC^{ht} = \sum_{p \in P} \sum_{l \in L} C_p^{ht} Cap_{l,p}^{ht} \quad (7.7)$$

POC grid electricity Cost

$$CC^{gridcost} = \sum_{l \in L} \sum_{t \in T} C_{l,t}^{FPVV} F_{l,t}^{e} \quad (7.8)$$

7.4 Constraints

Constraint expressions help to balance the flow within the overall system while accounting for the energy resources required, such as electricity flow or biomass resources for fuel-switching on heating technologies in industries. These are aided by flow terms with assigned indices within the sets which denote the flow between connections and components, which is important for, such as highly spatial model. In this section, system layers central to the study namely industry layer, biomass layer, generation layer and POC level are discussed. This section presents the core constraints required to interpret the results, while secondary formulations are provided in Appendix C.

7.4.1 Process Heat

The first series of equations is the heat energy supplied for the demands. This is to ensure all decarbonisable industrial process heat is fulfilled by a fuel technology. Equation (7.9) expresses the equation of heat load in factory sites $D_{l,t,q'}$ being supplied by heat flow from technologies $H_{l,t,p,q,q'}$.

$$D_{l,t,q'} = \sum_{p \in P} \sum_{q \in Q} H_{l,t,p,q,q'} \qquad \forall l \in L, t \in T, q' \in Q \quad (7.9)$$

The overall heat cascade can be done from a high-temperature band to a low-temperature band so a matrix $M_{p,q,q'}^{heat}$ is employed which is a parameter for feasibility.

$$\sum_{q' \in Q} H_{l,t,p,q,q'} = H_{l,t,p,q}^{sup} \quad (7.10)$$

$$H_{l,t,p,q,q'} \leq M_{p,q,q'}^{cascade} H_{l,t,p,q}^{sup} \quad (7.11)$$

As electric-based heating technology $p \in P^e$ can only produce a specific temperature q, a matrix $M_{p,q}^{e}$ is employed, whereas $M_{p,q}^{b}$ is included for the biomass energy term. However, $M_{p,q}^{b} = 1$ for all $p \in P^b$ and $q \in Q$. Equation (7.12) to Equation (7.14) describe the heat balance and generation from heating technologies.

$$H_{l,t,q}^{e,sup} = \sum_{p \in P^e} M_{p,q}^{e} F_{l,t,p}^{e,out} \quad (7.12)$$

$$H_{l,t,q}^{b,sup} = \sum_{p \in P^b} M_{p,q}^{b} F_{l,t,p}^{b,out} \quad (7.13)$$

$$H_{l,t,q}^{sup} = \sum_{p \in P^e} H_{l,t,p,q}^{e,sup} + \sum_{p \in P^b} H_{l,t,p,q}^{b,sup} \qquad \forall l, t, \quad (7.14)$$

7.4.2 Biomass heating technologies

The biomass variable for heating output can be represented as such for flow balance. η_p is added to account for the efficiency of the heating technology.

$$F_{l,t,p}^{b,out} = \sum_{q \in Q} F_{l,t,p,q}^{b,out} \qquad \forall l, t, p \in P^b \quad (7.15)$$

$$F_{l,t,p}^{b,out} = \eta_p F_{l,t,p}^{b,in} \qquad \forall l, t, p \in P^b \quad (7.16)$$

Heat technology sizing, including biomass technologies, was deliberately aligned with electricity-based technologies by associating capacity sizing with inlet energy flows rather than outlet energy. This decision is because outlet-based costing for heat pump technologies depends on the coefficient of performance, which varies with temperature lift. Using outlet energy as the sizing basis would therefore introduce inconsistency across temperature levels. The coefficient of oversizing heating technology $\alpha^{ht,sized}$ is fixed at 1.25 (Pag *et al.*, 2024).

$$\alpha^{ht,sized} F_{l,t,p}^{b,in} \leq Cap_{l,p}^{ht} \Delta t_t \qquad \forall p \in P^b \quad (7.17)$$

The mapping for suitable biomass feedstock type w is done for biomass heating technologies p.

$$\Omega^{bw} = \{(p, w) \in P^b \times W | p = \emptyset(w)\} \quad (7.18)$$

$$F_{l,t,p}^{b,in} = \sum_{m \in M} F_{l,t,p,m}^{b,in} \qquad \forall p \in P^b \quad (7.19)$$

7.4.3 Electric-driven heating technologies

Like the biomass efficiency, the energy transformation with thermal efficiency for electric-driven heating technologies η_p is represented by

$$F_{l,t,p}^{e,out} = \eta_p F_{l,t,p}^{e,in} \qquad \forall l \in L, t \in T, p \in P^e \quad (7.20)$$

Similar to the explanation for Equation (7.21). The expression for capacity for electric-driven heating technology is denoted by:

$$F_{l,t,p}^{e,in} \leq Cap_{l,p}^{ht} \Delta t_t \qquad \forall p \in P^e \tag{7.21}$$

The energy balance for electricity flowing into the factory from the POC layer for heating and electricity used for each heating technology. These terms can be expressed as:

$$F_{l,t}^{e} = \sum_{(i,l) \in A^{POC,fac}} F_{i,l,t}^{e} \tag{7.22}$$

$$F_{l,t}^{e} = \sum_{p \in P^e} F_{l,t,p}^{e,in} \tag{7.23}$$

7.4.4 Disjunction GDP Formulation in Transmission Lines

In order to avoid simultaneous bi-directional power exchange on the transmission link for each time step, a disjunctive constraint is introduced. For each pair of POC $\{i, u\}$ and time t, electricity can flow only in one direction. The constraint would prevent simultaneous opposing flows on the same transmission corridor, ultimately preventing artificial loop flows which could happen in a flow formulation.

$$\begin{bmatrix} F_{i,u,t}^{POC} \geq 0 \\ F_{u,i,t}^{POC} = 0 \end{bmatrix} \bigvee \begin{bmatrix} F_{i,u,t}^{POC} = 0 \\ F_{u,i,t}^{POC} \geq 0 \end{bmatrix} \tag{7.24}$$

The disjunction can be reformulated to a MILP format using a binary decision variable $z_{i,u,t}$. $z_{i,u,t}$ determines the flow direction for each connection for every time step t.

$$F_{i,u,t}^{POC} \leq \overline{F_{i,u,t}^{POC}} z_{i,u,t} \quad \forall \{i,u\}, t \tag{7.25}$$

$$F_{u,i,t}^{POC} \leq \overline{F_{i,u,t}^{POC}} \left(1 - z_{i,u,t}\right) \quad \forall \{i,u\}, t \tag{7.26}$$

$$z_{i,u,t} \in \{0,1\} \quad \forall \{i,u\}, t \tag{7.27}$$

Lastly, the POC nodal balance can be represented as:

$$\sum_{(j,i) \in A^{gen,POC}} F_{j,i,t}^{e} + \sum_{(u,i) \in A^{POC}} F_{u,i,t}^{POC} = D_{i,t}^{re} + \sum_{(i,l) \in A^{POC,fac}} F_{i,l,t}^{e} + \sum_{(i,u) \in A^{POC}} F_{i,u,t}^{POC} \quad \forall i \in I, t \in T. \tag{7.28}$$

7.5 Outputs and Reporting

7.5.1 Model Size and Computational Performance

From the P-graph representation framework, the model consists of 98,762 material nodes and 125,117 operating units. This creates a total of 340,959 directed edges between the nodes. Moreover, a total of 3,912 mutual exclusion constraints is enforced in the model to represent disjunctive operations and structural feasibility.

As the P-graph problem is translated into the Pyomo formulation, the original disjunctive programming formulation is reformulated into a MILP problem using the Big-M transformation method. The problem is then executed and solved using the Gurobi solver.

The final multiperiod optimisation model represents a large-scale mixed-integer programming problem. The complete formulation consists of 745,707 decision variables and 1,118,033 constraint equations, signifying a large and complex problem representing technology selection and structural disjunctions. The optimisation is executed on a 12th Gen Intel Core i7-12700H processor with 32 GB RAM of memory. The solver converged to an optimal solution with a MIP gap of 0.01.

Table 7.1 summarises the computational characteristics of the optimisation problem. The computational summary shows that a highly spatial multiperiod energy model can be solved within practical time limits while retaining detailed industrial, biomass and electricity network connections. The main bottleneck of the problem is defining the problem. The reason is because each set undergoes a rigorous search function so that the nodes specified in the P-graph problem do not overlap, and at the same time, it has searches in order to reduce the number of unnecessary variables and constraints which optimise the problem size.

Table 7.1: Summary of optimisation model size and computational performance.

Metrics	Values
Decision Variables	745,707
Constraints	1,118,033
Binary Variables	315,786
Continuous Variables	429,921
Time for defining the problem (s)	4,264.2
Time for building the problem (s)	711.3
Time for solving the problem (s)	553.0
Total elapse time (s)	6,319.0

Output Structure and Discussion

The optimisation produces detailed outputs across multiple interconnected layers of the industrial energy system. At the site level, the model determines optimal fuel-switching configurations, including fully electrified, fully biomass-based, and hybrid arrangements combining both technologies. The key outputs include site-specific heating technology, energy

flow across each industrial site, electricity consumption by site and time period, biomass consumption by site and time period and biomass transport flows between nodes. These outputs are to be then post-processed and aggregated by region, sector and fuel type.

Regional Patterns

The regional distribution of decarbonised process heat reveals clear spatial differentiation as displayed in Figure 7.2. Biomass-based heating dominates in biomass-rich regions, particularly Northland, the lower and central North Island, and large parts of the South Island. This outcome is structurally driven by proximity to biomass resources and lower associated transport costs. The optimisation internalises transport distance and resource availability, which naturally favours biomass deployment in these regions.

The regional breakdown (Figure 7.2) further illustrates that electrification and biomass are not uniformly distributed across the country, but instead reflect spatial heterogeneity in resource availability, industrial composition, and process temperature requirements.

Electrification is more prevalent in urbanised and semi-urban regions such as Auckland, Bay of Plenty, Otago, and Canterbury. Urbanisation and limited local biomass availability, along with lower temperature process heat requirements, specifically lead to higher electrification pathways in Auckland (Figure 7.3). A closer examination of the Auckland region indicates that many sites operate within low to medium-temperature ranges, making electrification via heat pumps technically efficient and economically attractive under the model assumptions.

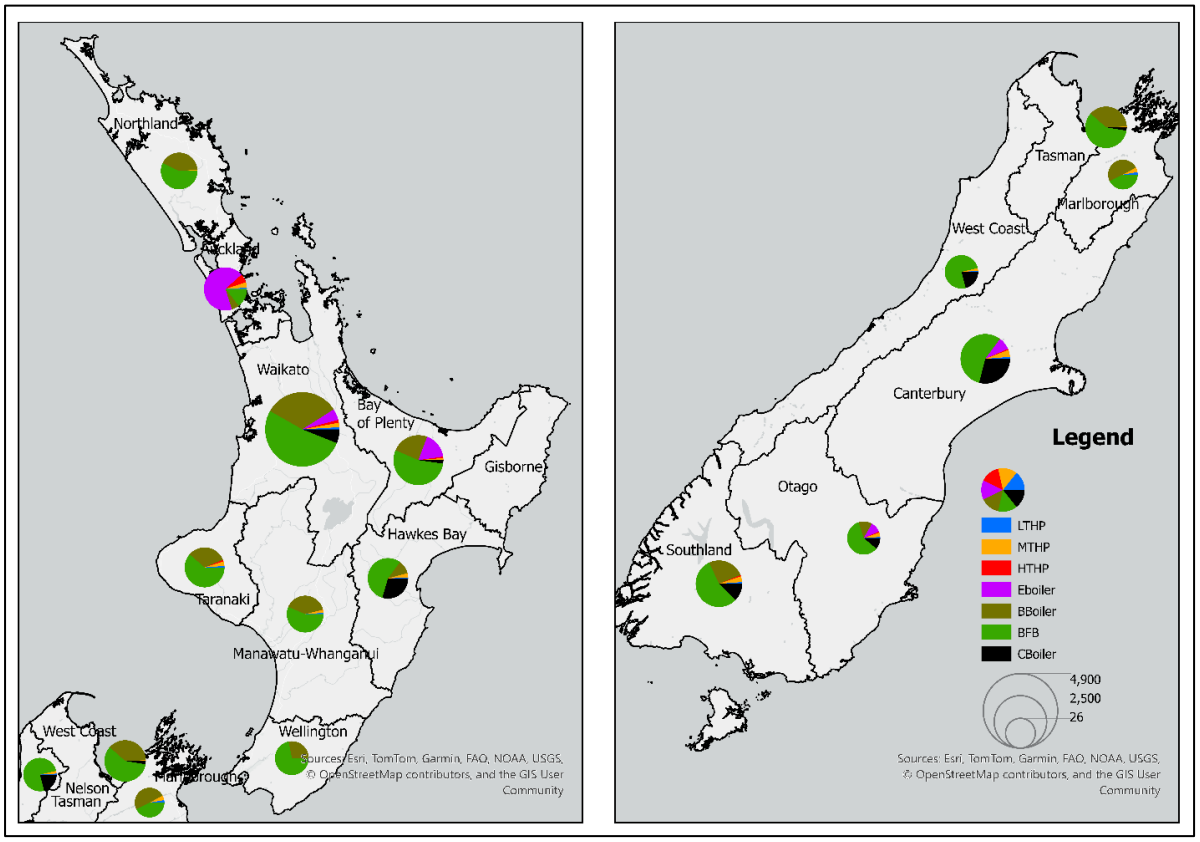


Figure 7.2: Regional distribution of decarbonisable process heat by input fuel to heating technologies.

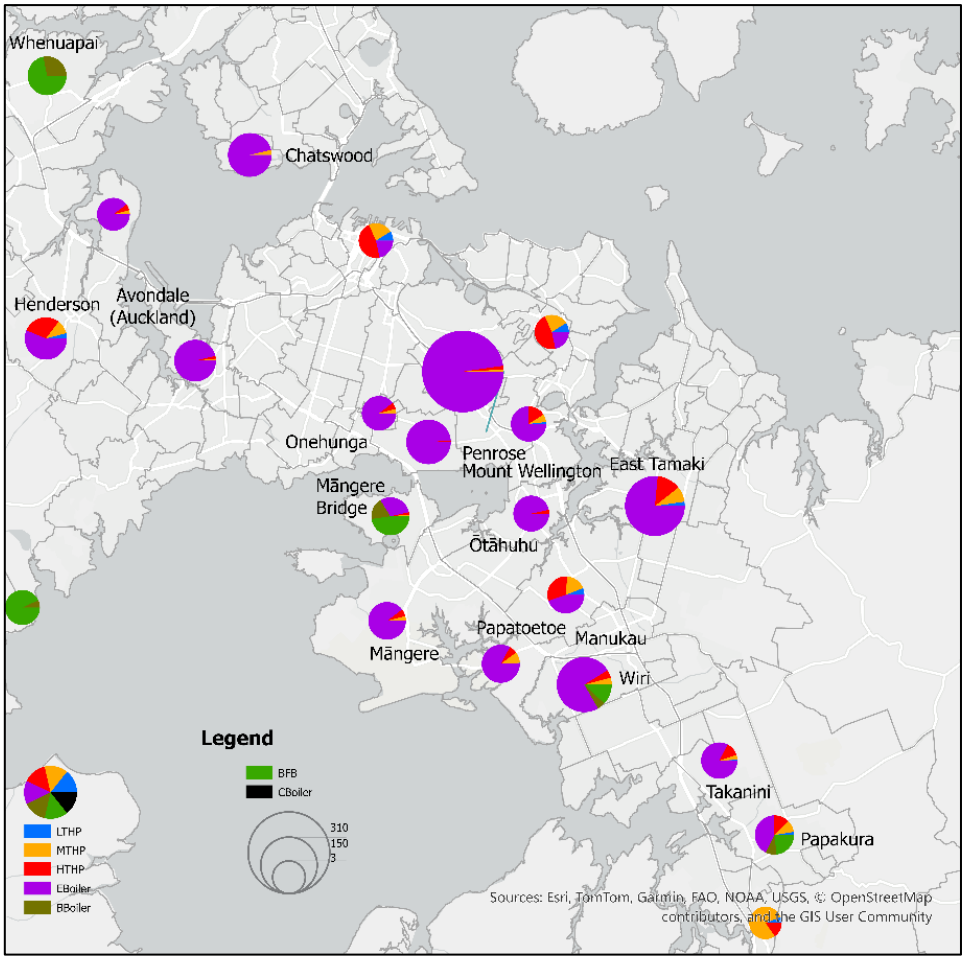


Figure 7.3: Fuel-switching technologies by input fuel in the Auckland region.

Sectoral Patterns

Figure 7.4 presents fuel intake by industrial sector. The results show strong sectoral differentiation. The process heat in wood processing and pulp and paper industries is almost exclusively biomass-based. This outcome reflects both structural and economic factors. These industries are typically located near forestry resources and often generate internal biomass residues, which reduces effective fuel cost and transport dependency. Conversely, the food processing sector exhibits a high degree of electrification, whereas dairy processing shows a more mixed configuration between electrification and biomass. The difference is largely driven by temperature requirements. Low-temperature process heat is frequently electrified using heat pumps due to favourable thermodynamic performance, leading to achieving high COP and cost competitiveness. Higher temperature applications remain more reliant on biomass or hybrid configurations, as electrification at higher temperatures becomes more capital-intensive.

After adjustments with technology efficiency as shown in Figure 7.5, the end-use demand-fulfilling industries show a more important role of electrification in supplying heat to the factory system. This change is largely contributed by the high energy output from the low-temperature and medium-temperature heat pumps as they have COPs of 4.57 and 2.69 due to the small

temperature lift. The model shows a share of electrification based on end-use demand to be 4.32 TWh yr^{-1}, significantly higher than when quantifying it with fuel intake, which is 2.75 TWh yr^{-1}.

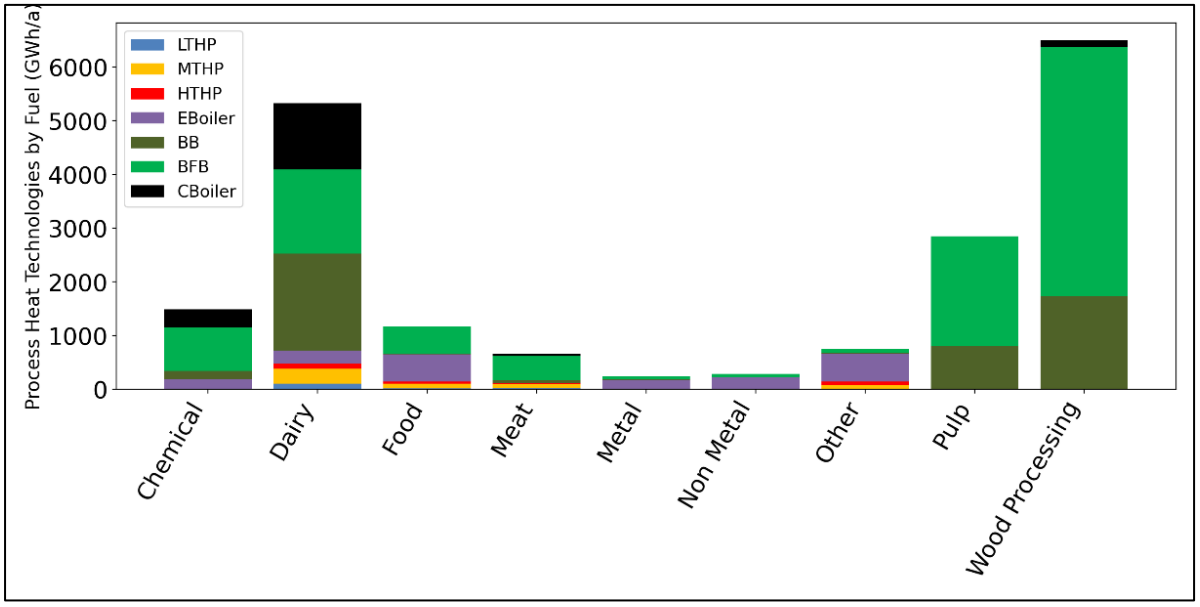


Figure 7.4: Decarbonisable process heat by sector and input fuel for heating technologies.

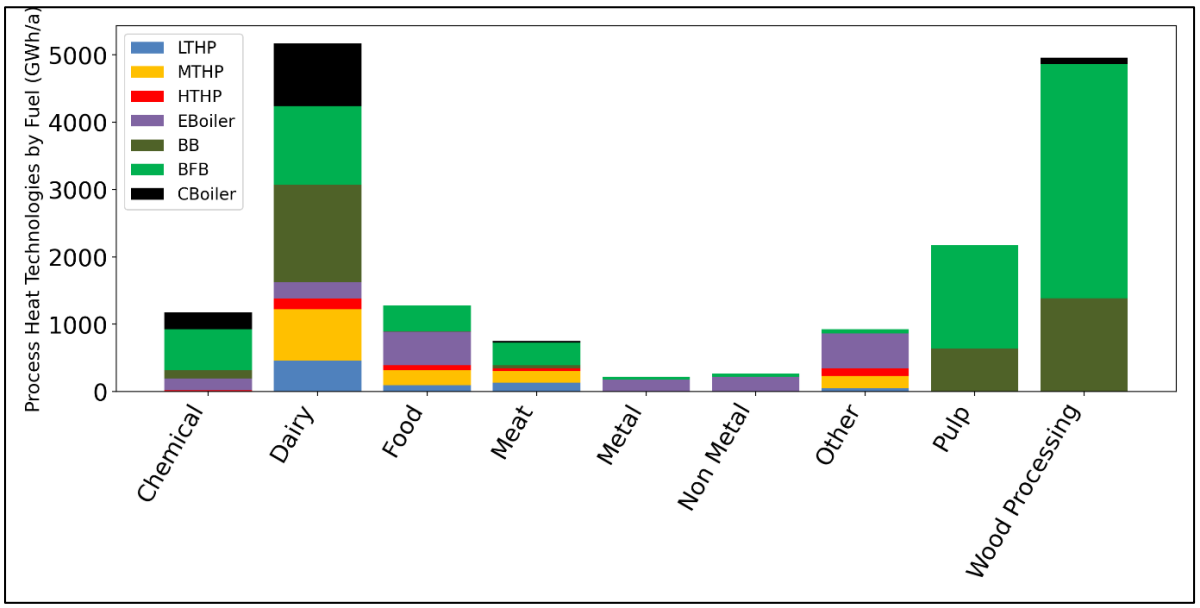


Figure 7.5: Decarbonisable process heat by sector with end-use demand for heating technologies.

These sectoral patterns demonstrate that technology selection is not uniform across industries but rather depends on temperature demand profiles, geographic location, and resource accessibility.

Coal Boiler Retrofitting

Of the identified coal boiler capacity of 828 MW remaining after filtering by age and size constraints, approximately 456 MW is economically retrofitted to biomass under the least-cost optimisation. This is further calculated 1,743.1 GWh per annum of woody biomass fuel is supplied

to coal-retrofitted boilers for heat supply. This indicates that retrofitting existing coal boilers to biomass remains a structurally competitive decarbonisation pathway under current fuel cost and resource availability assumptions.

Figure 7.6 illustrates the spatial distribution of retrofitted biomass capacities. The results indicate that the majority of viable coal boilers are converted, as biomass retrofitting presents a relatively low-cost transition pathway compared to full electrification, particularly for medium to high-temperature applications. This finding reinforces the importance of transitional retrofit strategies in achieving near-term emissions reductions.

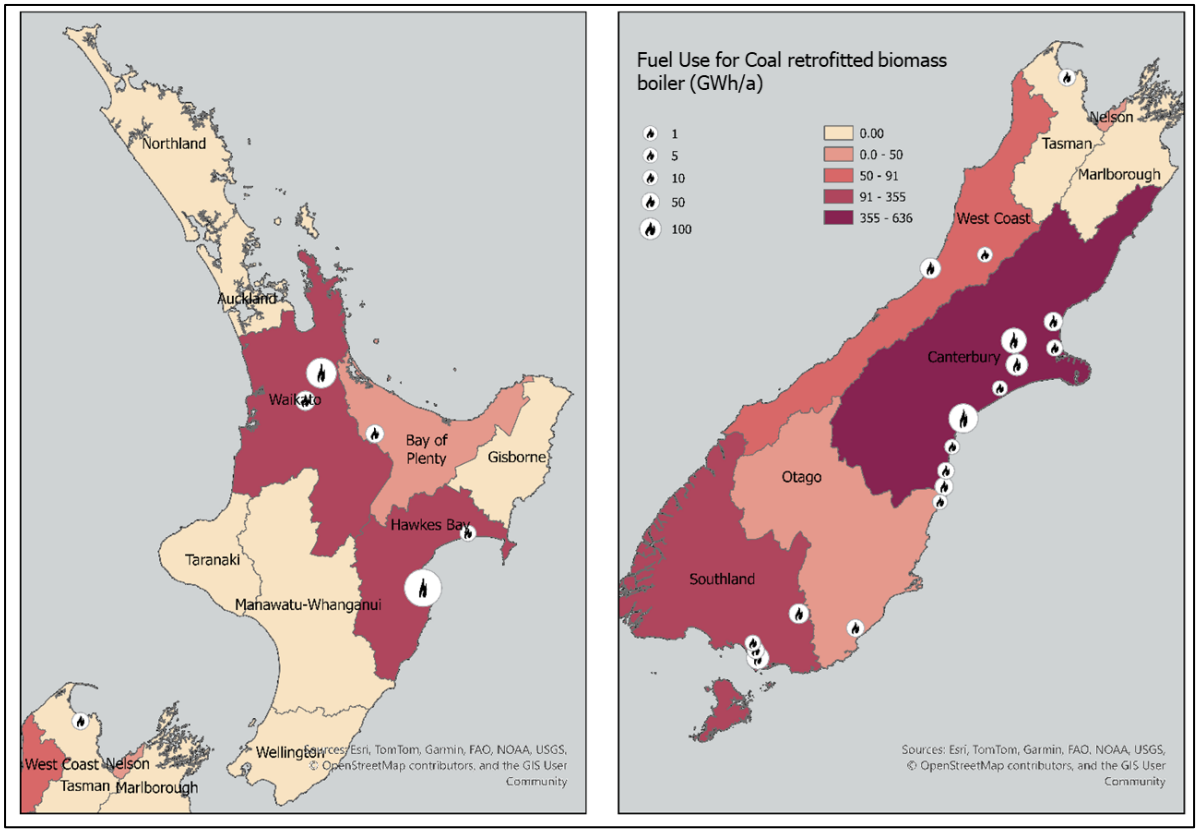


Figure 7.6: The spatial distribution of retrofitted coal boilers.

Biomass Type Utilisation and Transport Structure.

The model differentiates between biomass types, including woody biomass and non-woody alternatives such as straw and stover. Woody biomass is selected more frequently across regions. This outcome reflects differences in effective cost per unit of energy, transport feasibility and more fuel-efficient for producing heat. The optimisation establishes 559 active biomass transport connections between resource nodes and industrial facilities, as displayed in Figure 7.7. When aggregated annually, these flows reveal concentrated biomass transport corridors in regions with dense industrial clusters. Transport intensity is highest where biomass availability and industrial heat demand co-locate, reinforcing the importance of regional resource planning.

The transport structure also highlights the endogenous balancing of biomass allocation across competing industrial uses. Rather than allocating biomass uniformly, the model directs flows

toward locations where the marginal decarbonisation benefit is highest relative to cost. Based on the results, the optimised model shows that biomass transport tends to favour shorter travel distances before selecting long distance travels. Figure 7.8 shows that the modal transport distance falls within the 10–50 km range for biomass transport.

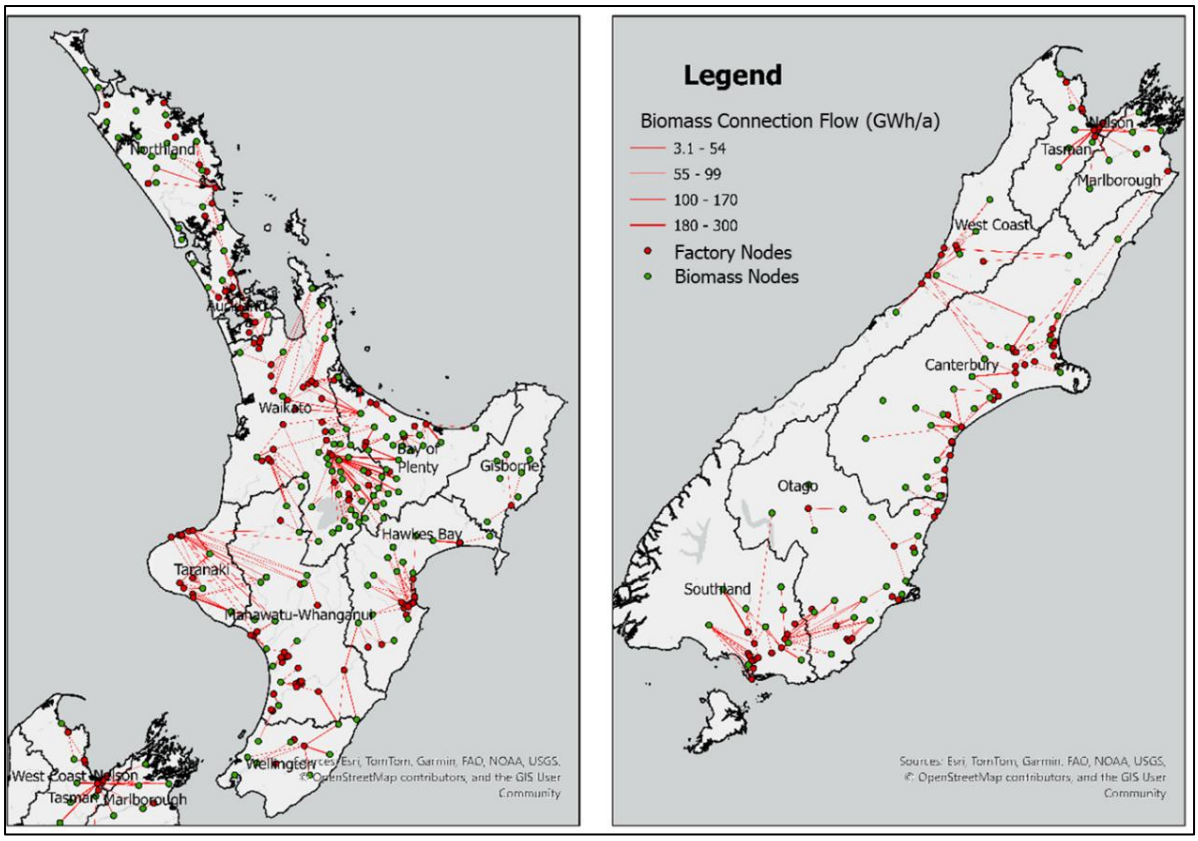


Figure 7.7: The spatial distribution of the biomass connections to factories.

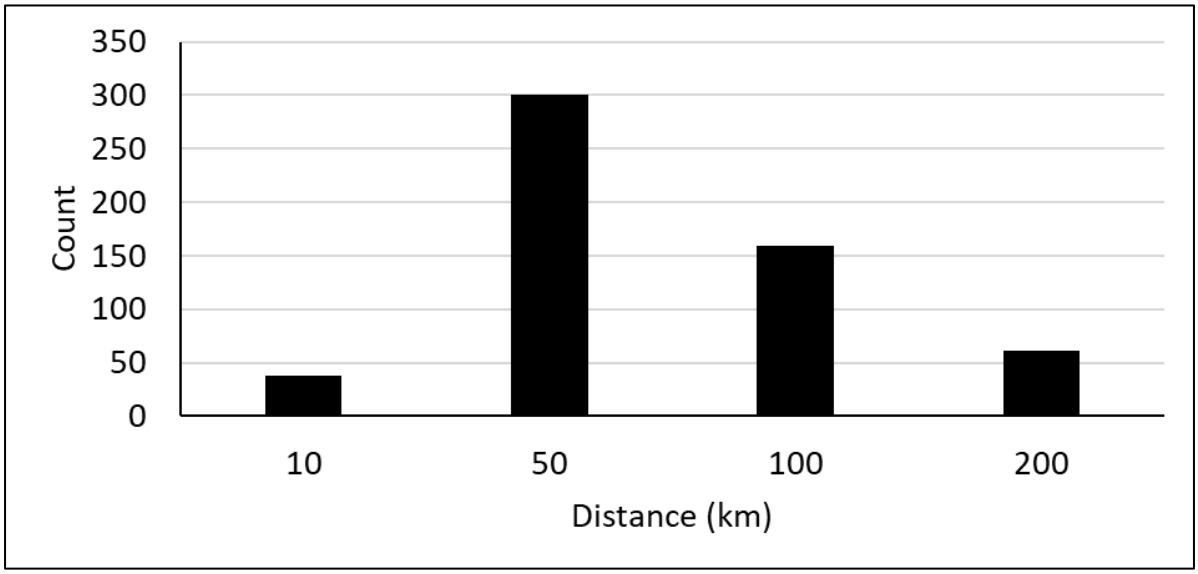


Figure 7.8: Number of biomass to factory connections established in the output results.

On the other hand, biomass utilisation remains high owing to a large amount of biomass consumed in the solution results. Only certain locations, such as Taranaki and the West Coast

region, show saturated biomass utilisation, which indicates that overall, there is a moderate availability of biomass for factory heat consumption with the condition that there are no competing uses for these resources. The biomass recoverability level used in the base scenario is L0, which represents the midpoint between the high recoverability (L1) and low recoverability (L2) scenarios. Figure 7.9 and Figure 7.10 outline the biomass resource use by region.

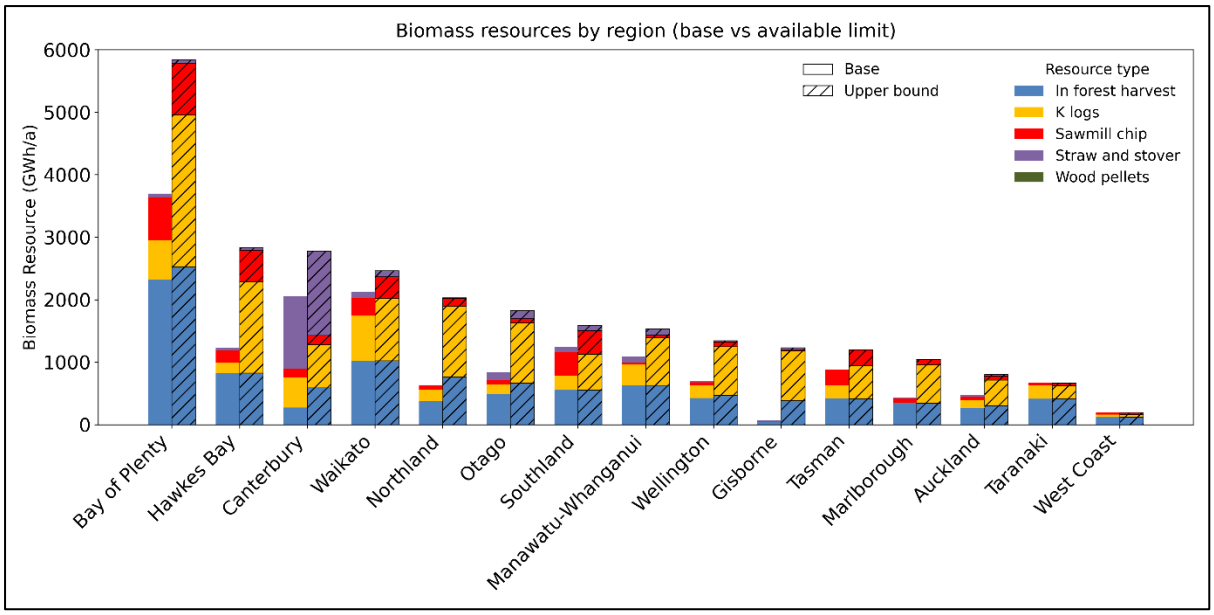


Figure 7.9: Biomass resources utilised and upper limit by region.

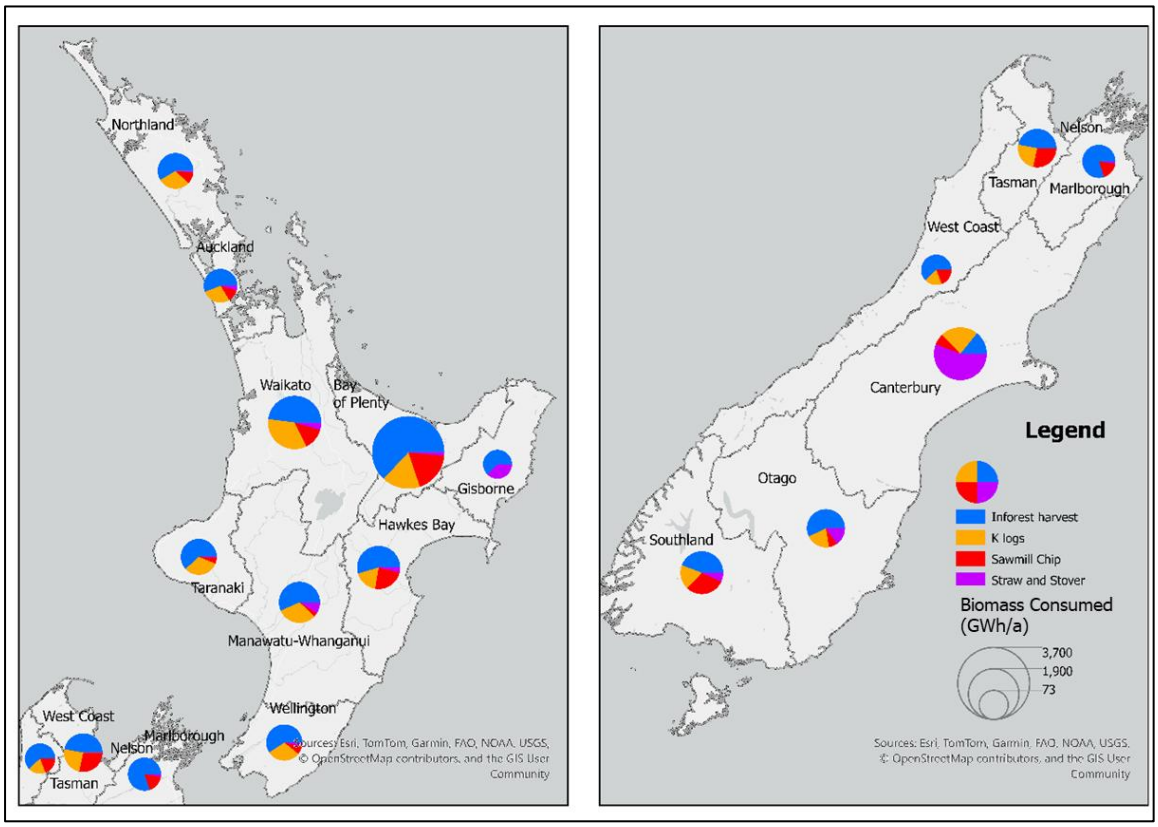


Figure 7.10: Figure of biomass type consumed by region for process heat industry.

Owing to the high granularity of the energy model, the model detail can go down to the nodal level. Figure 7.11 shows the connection pathways from an individual biomass node. The exemplary node is further supported by the biomass flow and resource type to corresponding sites as shown in Figure 7.12.

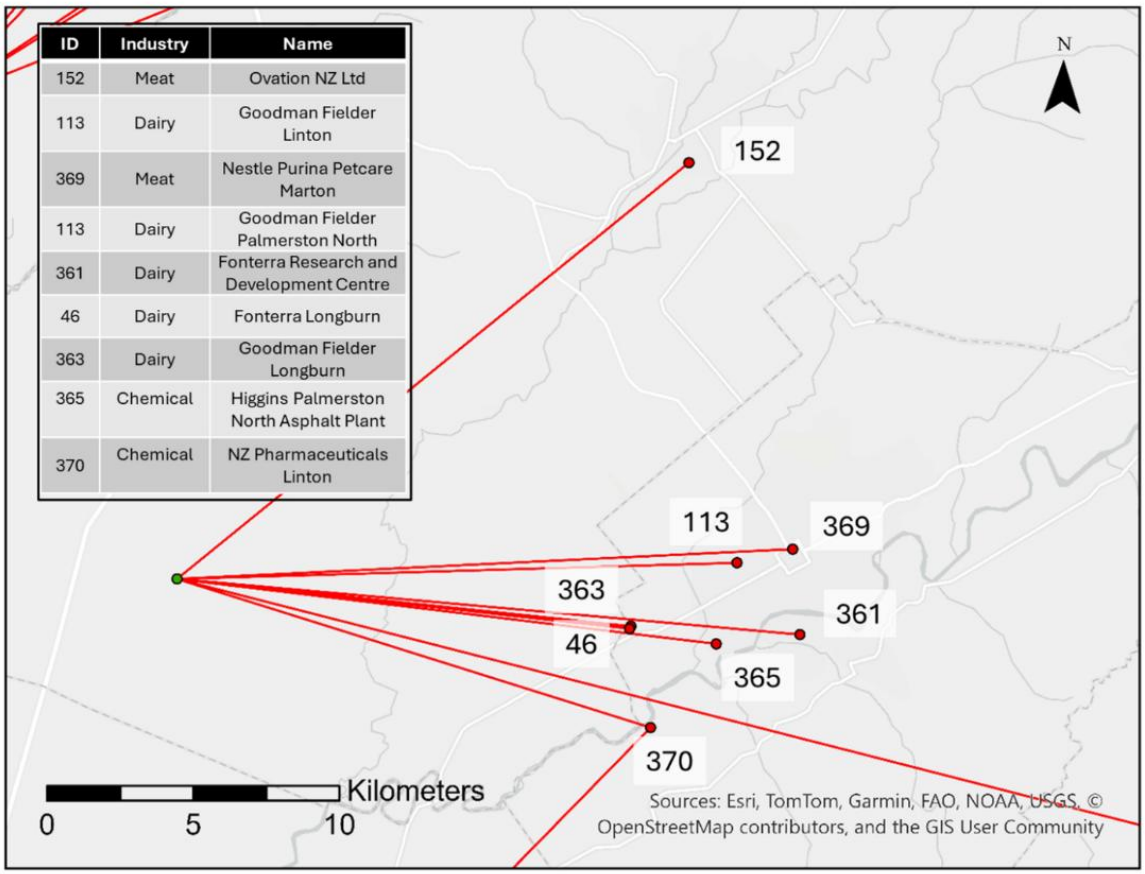


Figure 7.11: Example biomass source (ID 81) and its connections.

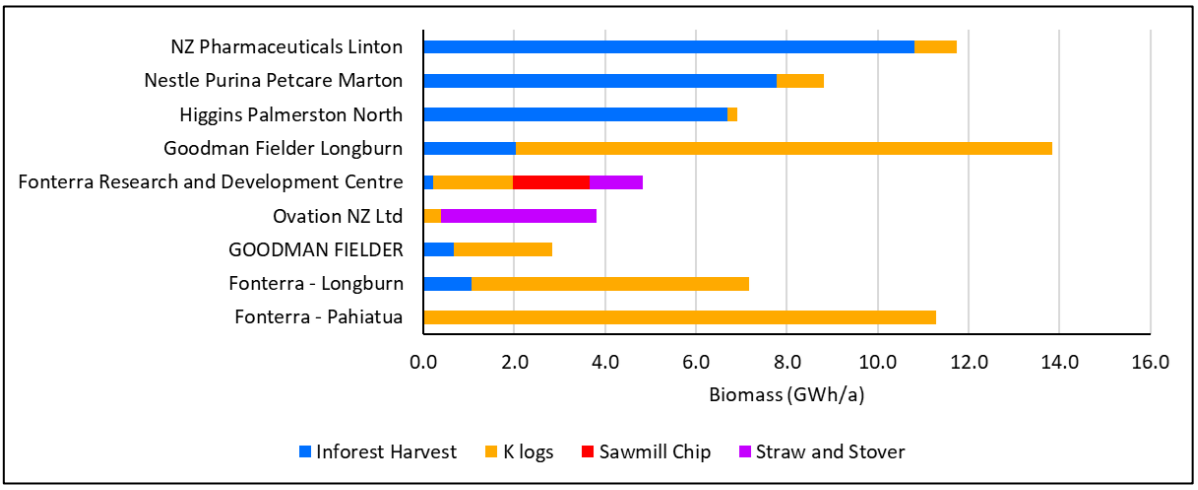


Figure 7.12: Total biomass flow from ID 81 to each destination ID aggregated over the full year as the sum of monthly flows.

Green Methanol

Green methanol is only profitable when the green methanol price is set at 870 NZD t^{-1} (Figure 7.13). This output revealed that green methanol is not economically feasible with the current production pathway where electrolysed hydrogen reacts with carbon from biomass for the production of green methanol. Only several locations have high green methanol potential where biomass resource is abundant, as it is a competing resource between green methanol and process heating. This shows that green methanol is not economically feasible considering the selling price is 600 NZD t^{-1} for 2024. Figure 7.14 provides a spatial overview of potential feasible green methanol locations when the selling price is set at 950 NZD t^{-1}.

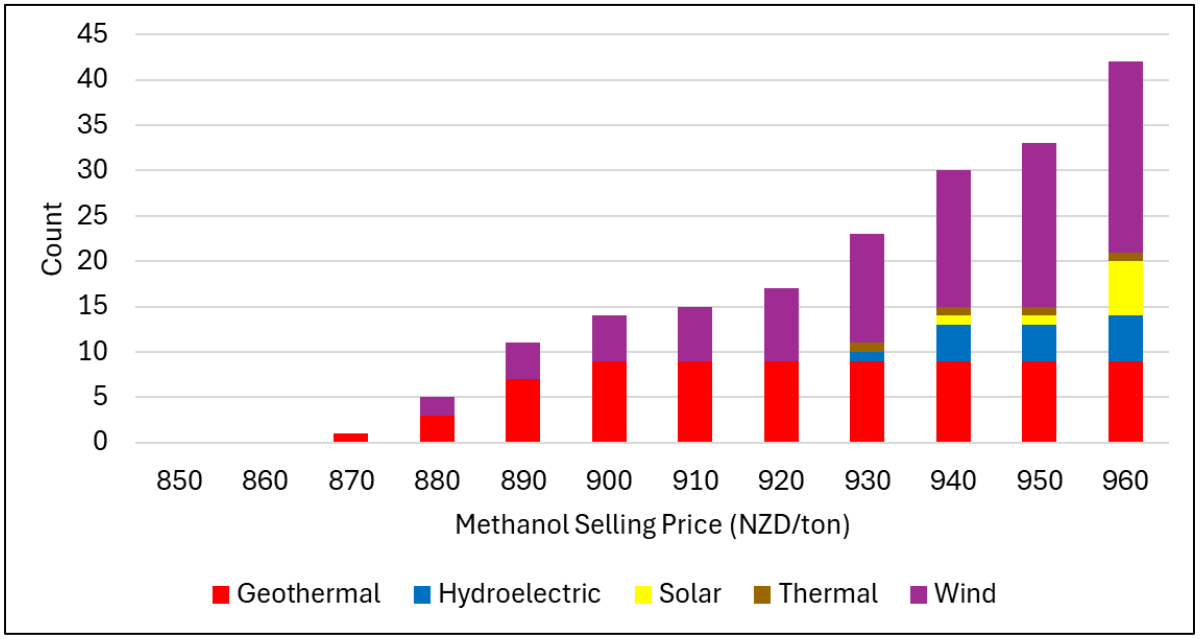


Figure 7.13: Number of green methanol feasibility sites by varying selling price.

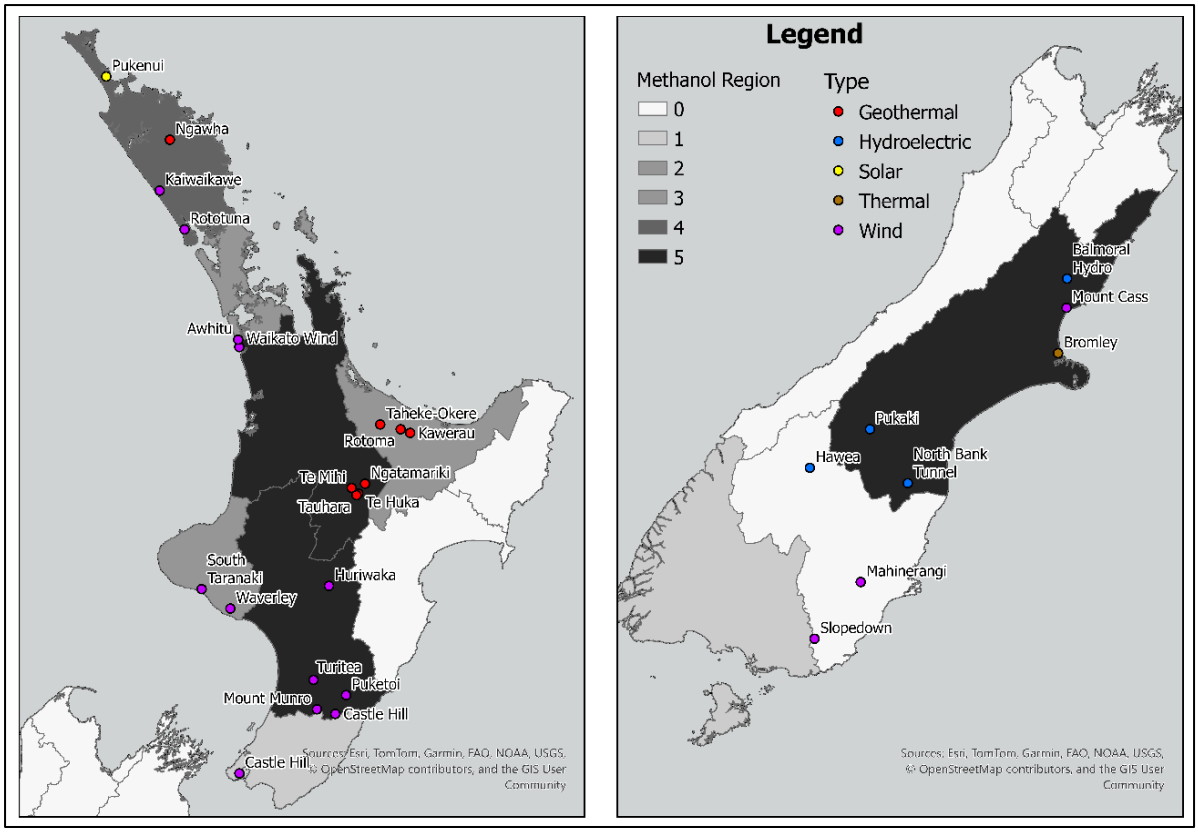


Figure 7.14: Possible green methanol sites when selling price is set as NZD 950 NZD t^{-1}.

7.5.2 Model Verification and Consistency Checks

As an optimisation-based energy system model, a direct experimental validation is not possible. Instead, model credibility is assessed through a series of consistency checks by comparing with decarbonisation initiative projects announced by sites on fuel-switching proposals. As the scale of the dataset includes 298 industrial sites, it is not feasible to validate each industrial site in detail. Hence, an approach involving selecting several publicly disclosed decarbonisation plans to validate the energy modelling results to align with real-world investment decisions.

Comparison Against Announced Industrial Projects

Fonterra sites are selected as primary validation cases due to the transparency of their published decarbonisation strategies and their relevance to national emissions reduction targets. Moreover, the sites are scattered all across the country to show the effect of spatial attributes which influence the outcomes of the fuel-switching technologies, such as variations in local energy sources, infrastructure availability, and regional policies that may affect implementation.

The optimisation results are compared qualitatively against the decarbonisation plans. The model suggests biomass conversion in biomass-rich regions and electrification or electrode boiler deployment in regions with favourable electricity access and suitable temperature profiles. These outcomes are fairly consistent with the decisions made by sites. In addition, an additional scenario

is developed to examine the point when fuel-switching changes, specifically when biomass-based heating technologies shift toward electricity-dominated heating. Hence, results from a scenario where biomass cost is increased by 300% are included in the model in addition to the results from the base scenario. This scenario is developed under the assumption that biomass prices may increase due to rising market demand for biomass combustion. Table 7.2 summarises publicly announced projects and their commissioning timelines.

Table 7.2: Summary of company sites identified for comparison.

Site	Announced Plan	Commission Date	Source
Fonterra Stirling	Biomass Boiler	August 2023	Fonterra New Zealand (2022)
Fonterra Whareroa	Heat Pumps	September 2023	Fonterra New Zealand (2022)
Fonterra Clandeboye	Biomass boilers	February 2024	Fonterra New Zealand (2022)
Fonterra Hautapu	Boiler Conversion	August 2024	Fonterra New Zealand (2022)
Fonterra Waitoa	Biomass Boiler	October 2024	Fonterra New Zealand (2022)
Fonterra Edendale	Electrode Boiler	October 2024	Fonterra New Zealand (2022)
Goodman Fielder Christchurch	Heat Pumps	September 2025	EECA (2025a)
New Zealand Sugar	MVR	September 2025	(2025a)
Open Dairy Country Waharoa	Biomass Boiler	January 2024	Open Country Dairy (2026)

This alignment provides external validation that the model captures realistic economic and technological drivers influencing industrial decarbonisation decisions. Hence, the following section presents the end-use demand distribution across heating technologies for process heat.

Fonterra Stirling

For the Fonterra Stirling plant, the site has announced the project to phase out coal boilers with biomass boilers. The optimisation model for the base scenario suggests the application of the use of coal-retrofitted biomass boilers as the primary heat source for the site. On the other hand, biomass is rather dominated despite having biomass prices increase drastically. This shows that the decision for Fonterra Stirling has a high threshold before biomass-based heating shifts to electricity.

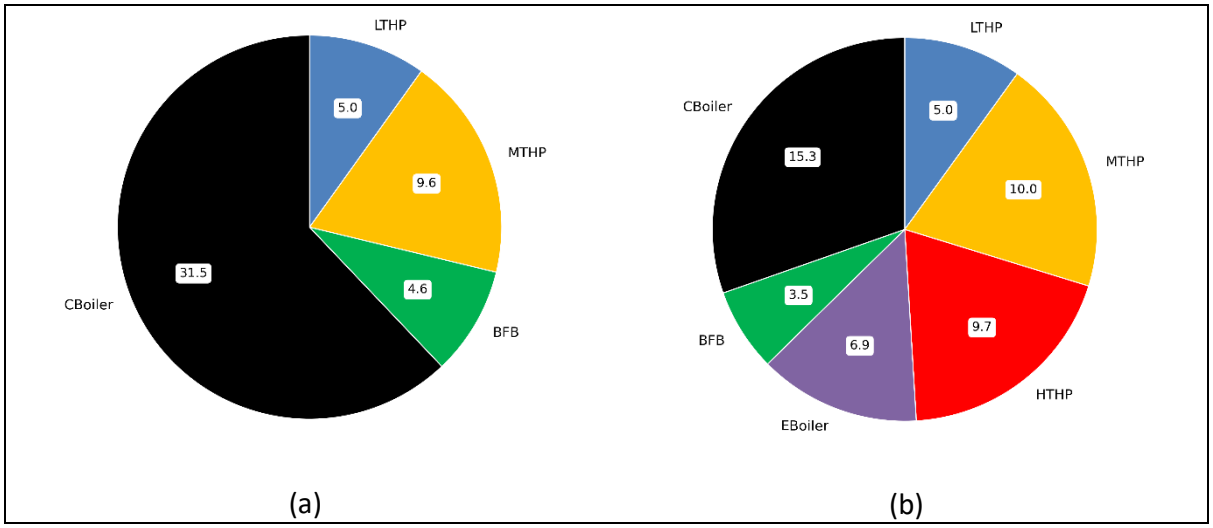


Figure 7.15: Model results in GWh yr^{-1} for the Fonterra Stirling site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%.

Fonterra Whareroa

For the Whareroa heat pumps, there is a fair presence of heat pump technologies proposed by the model. It is aligned with the decarbonisation of replacing 9,100 tonnes of carbon dioxide (approximately 3,300 tonnes or 39.6 GWh of natural gas equivalent per annum). In Fonterra Whareroa, the roles of HTHP and EB replace the biomass boiler (BB) once biomass prices increase significantly. However, the deployment of BFB technology still remains, demonstrating the resilience of non-woody biomass resources in supporting process heat supply.

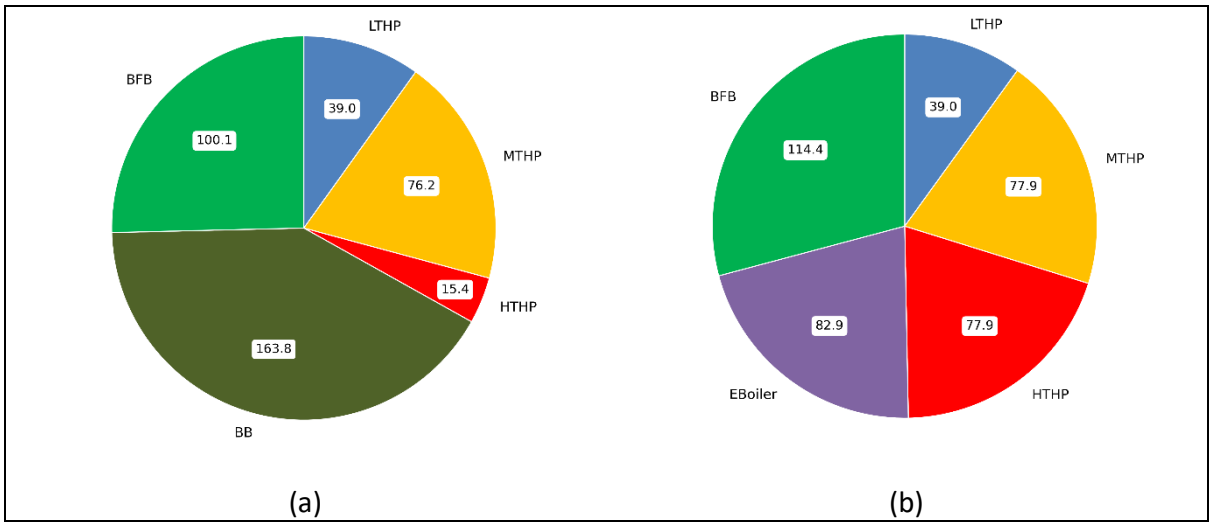


Figure 7.16: Model results in GWh yr^{-1} for the Fonterra Whareroa site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%.

Fonterra Clandeboye

Foneterra Clandeboye has conducted heat recovery projects and biomass to convert their two coal boilers to wood pellets, which is crucial for their phasing out of coal as a heating resource by

2037. The model suggested a similar path with recommendations, which is to retrofit Clandeboye boilers for biomass use. As the price of biomass becomes relatively high in the biomass scenario, HTHP plays a role in reducing the amount of biomass needed.

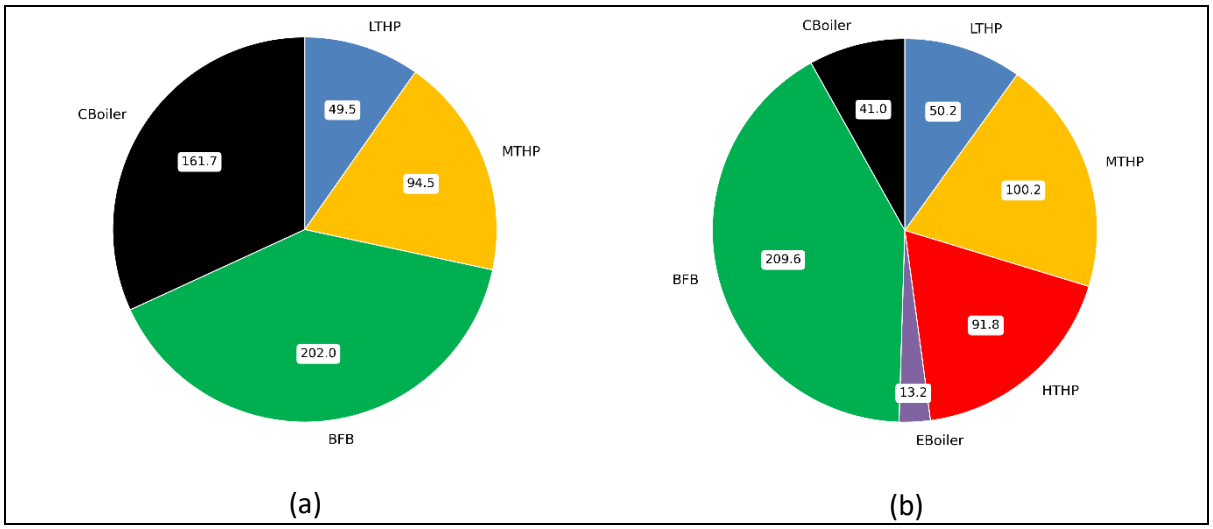


Figure 7.17: Model results in GWh yr^{-1} for the Fonterra Clandeboye site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%.

Fonterra Hautapu

Similar to Clandeboye, Hautapu would undergo boiler conversion to reduce 12000 tonnes of CO_2 or 5200 tonnes or 33.9 GWh of coal as part of their decarbonisation effort. The energy model shows that the majority of decarbonised heat can be supplied by a retrofitted boiler, with a supply of heat pumps as well. Instead of retrofitting coal boilers for woody biomass use, the results show that a hybrid configuration of EB and BFB technologies is deployed in the second scenario, making boiler conversion unnecessary, which suggests that this approach may be more efficient and environmentally friendly than traditional methods.

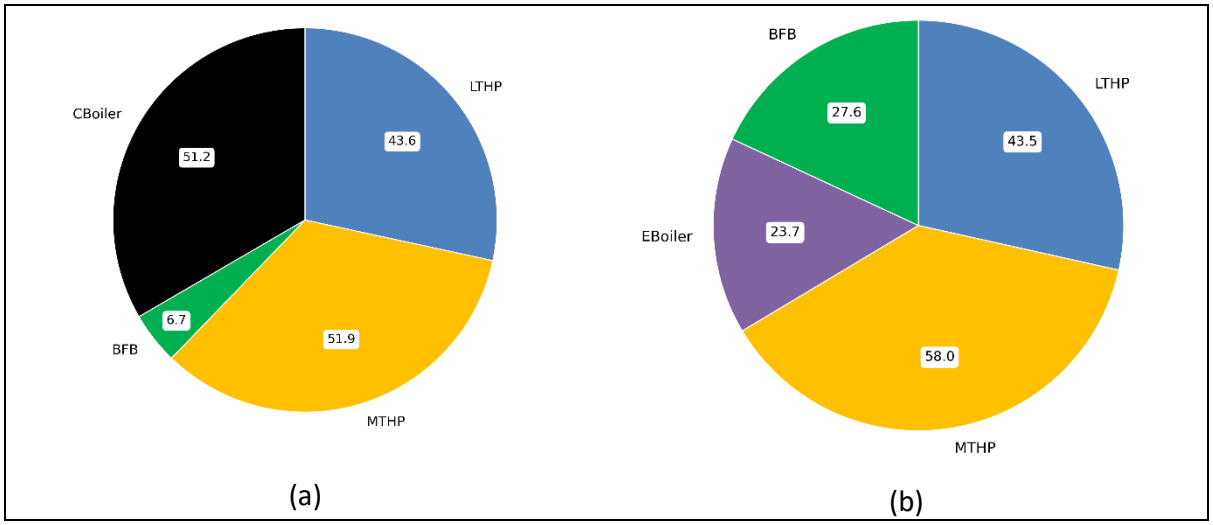


Figure 7.18: Model results in GWh yr^{-1} for the Fonterra Hautapu site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 400%.

Fonterra Waitoa

Waitoa has also halved the use of coal on-site with the introduction of new boilers, reducing its emissions by 46000 tonnes of CO_2-equivalent or 20000 tonnes of coal, or 130 GWh of thermal energy. Although Waitoa decided to introduce new biomass boilers instead, the optimisation model suggests repurposing the boilers for further cost reduction purposes. Nevertheless, the fuel type used in the model remained the same, which is biomass. As coal boilers are modular, the model favours flexibly sizing new EB and HTHP technologies under high biomass price conditions. These technologies help mitigate the impact of rising biomass prices.

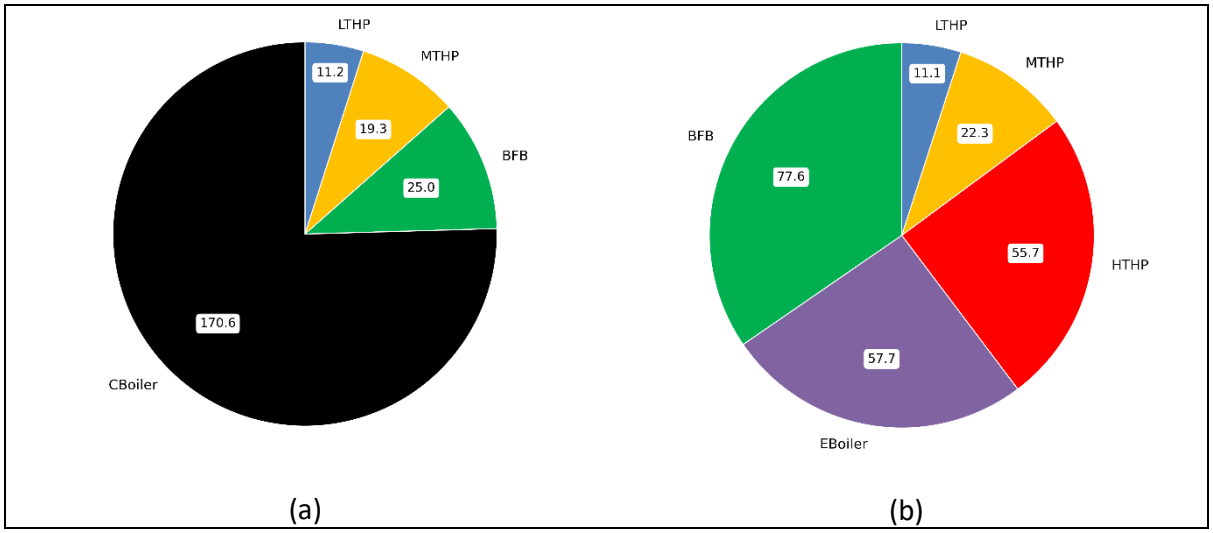


Figure 7.19: Model results in GWh yr^{-1} for the Fonterra Waitoa site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%.

Fonterra Edendale

Despite Fonterra's decision to electrify the site, the optimisation model suggests the opposite. The energy model recommends using biomass bubbling fluidised bed boilers. As the model suggests the biomass can be supplied from forest sites (ID=1, 4, 7) which are 26.4 km, 9.8 km and 30.5 km away from the site. Part of the reason could be the unknown electricity negotiation price contracts between Edendale and Meridian Energy, which would be significantly lower than the estimated electricity prices which are made across all POCs.

To ascertain the electricity price for the site to be electrified, the sales electricity price of the site is modified to half to assess the electricity price and double the biomass where the model would swap from the biomass share to the electrification share. From the modelling results, it can be estimated that the electricity price negotiated by the Edendale site would need to be at least half of the average POC adjusted electricity price in order to make electrification a viable fuel-switching decision.

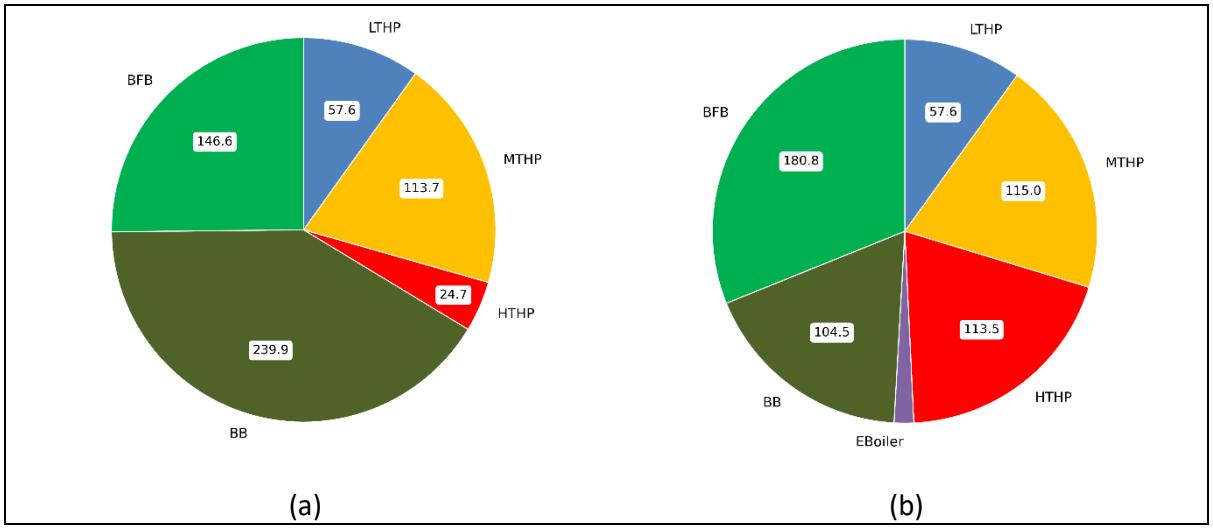


Figure 7.20: Model results in GWh yr^{-1} for the Fonterra Edendale site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 100% and electricity price for the Edendale site is halved.

Goodman Fielder Christchurch

Goodman Fielder deployed industrial heat pumps to capture heat from the ammonia refrigeration system, where the site claimed this would reduce 35% of carbon emissions at the site. Two 450 kW heat pumps were deployed to replace gas boilers in the heating system. The model also suggests a good amount of high and medium-temperature heat pumps, which would generate 7.37 GWh of thermal energy from 3.73 GWh of electricity. This site is purely electrical due to the site being located in an urban area as defined by Stats NZ (2024).

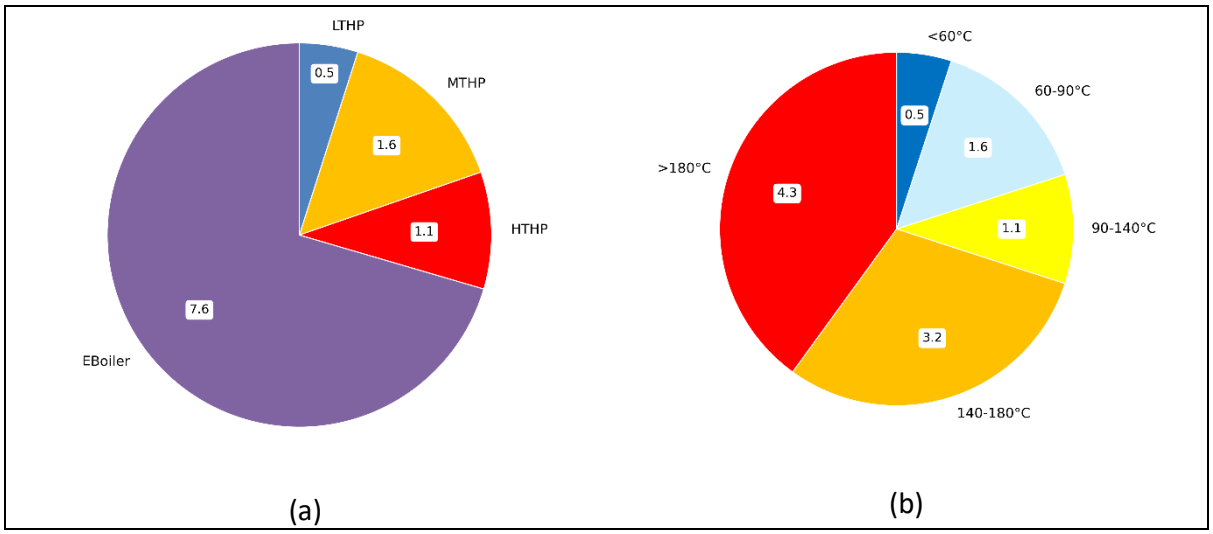


Figure 7.21: (a) Optimisation model results in GWh yr^{-1} for the Fonterra Edendale site under the base scenario; (b) temperature-based demand requirements of the main industrial processes at the site.

New Zealand Sugar

As an example, New Zealand Sugar is replacing its old steam-driven thermal vapour recompression evaporator with an electric mechanical vapour recompression (MVR) system. However, as the MVR system which is an indirect heating system, does not explicitly exist in the model and the temperature range falls under a boiler category in the model, it is represented differently. However, the model recommends electrification for this temperature range, indicating that New Zealand Sugar's plan aligns with the model results. Similar to the Goodman Fielder Christchurch site, New Zealand Sugar is also in an urban area, implying that electrification is needed for decarbonisation.

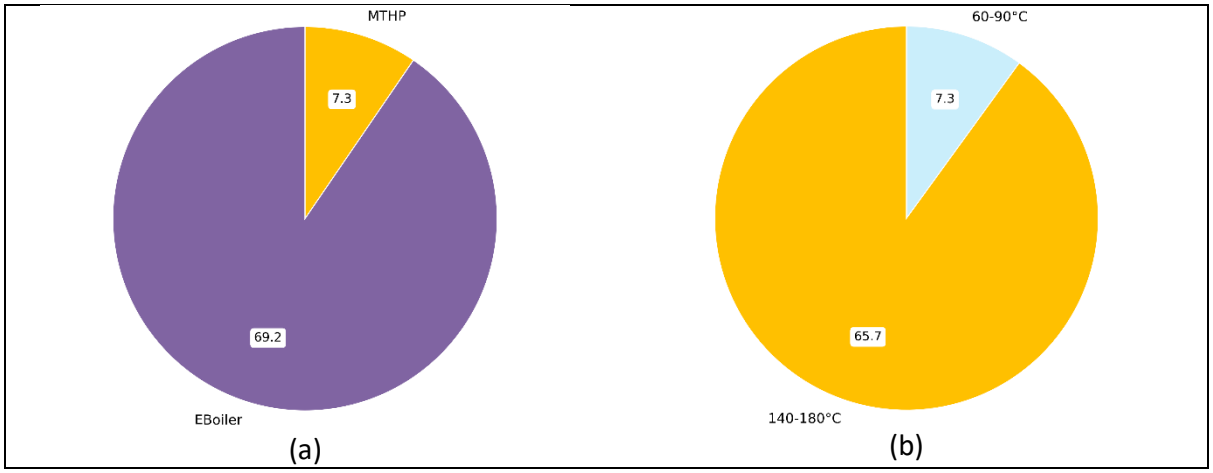


Figure 7.22: (a) Optimisation model results in GWh yr^{-1} for the NZ Sugar site under the base scenario; (b) temperature-based demand requirements of the main industrial processes at the site.

Open Dairy Country Waharoa

The decarbonisation project in Waharoa has been commissioned to switch coal to wood pellets at the site. The project is designed to reduce greenhouse gas emissions by over 70%. The model suggests that almost 97.9% of biomass is used to replace coal as the primary heat energy, where 80.8% is heated using a biomass boiler and the remaining is heated applying a bubbling fluidised bed boiler. This shows a strong presence relation of biomass presence in the site for fuel-switching.

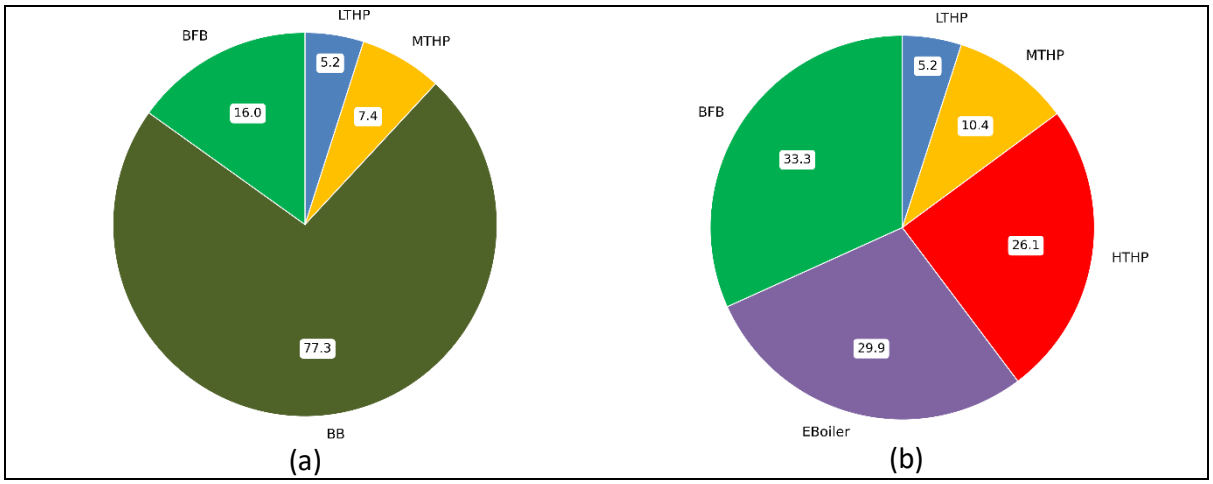


Figure 7.23: Model results in GWh yr^{-1} for the Open Country Waharoa site under different biomass cost assumptions: (a) base scenario; (b) scenario where biomass cost is increased by 300%.

7.6 Sensitivity and Robustness Checks

Sensitivity analysis is conducted on key parameters which are studied to have the highest degree of uncertainty which would affect the share of the outcome on fuel-switching in technologies. These include electricity price, biomass cost, and recoverability factors. The electricity price of -50% and +100% of the base value is considered in order to estimate uncertainties such as a dry year which would potentially be a concern in the future due to climate change. This is because the industrial electricity price is not consistent, and electricity prices would shoot up during dry years.

On the other hand, biomass resources of L1 and L2 define the high estimate and low estimate of the different recoverability rates from the gross biomass resource scenarios made by the Bioenergy Association of New Zealand as displayed in Table 7.3. A new L3 scenario is added for the sensitivity analysis to assess a scenario where there is lower availability of resources as well as competing with other uses. Sawmill chip is reduced to 0.1, as it is an intermediate product in the wood processing industry.

Besides that, as there will be a certain variability of the biomass price due to the energy content and the market fluctuation, the Bioenergy Association also concluded that there is a degree of price variability due to the difference in moisture level and cost of extracting biomass resources based on locality and availability (Bioenergy Association of New Zealand, 2021). In the sensitivity analysis model, the biomass price is varied from -50% to 400%. Figure 7.24 shows the electrification rate of industrial process heat, excluding the wood processing and pulp and paper sectors.

Table 7.3: Recoverability factors based on different estimations by Scion for sensitivity analysis (Hall, 2022).

Biomass Type	L1	L2	L3
In-forest harvest	0.8	0.65	0.5
K logs	0.95	0.8	0.6
Sawmill Chip	0.7	0.6	0.1
Straw and Stover	0.75	0.5	0.4
Pellets	1	1	1

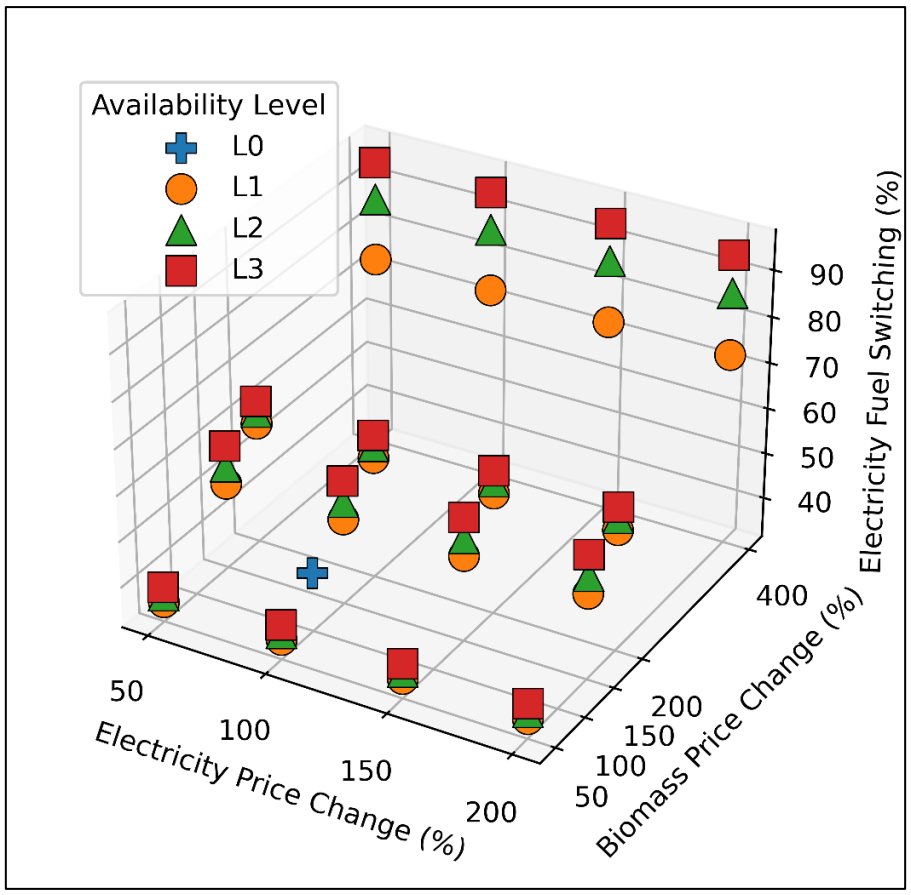


Figure 7.24: Sensitivity analysis of electrification uptake in industrial process heating (excluding wood processing and pulp and paper sectors).

The direction of technology switching under parameter variation remains economically intuitive and further supports internal model consistency. The trends are as follows:

- Higher electricity prices shift solutions toward biomass.
- Lower biomass availability increases electrification.
- Increased biomass resource cost reduces biomass use as a heating fuel.

Based on the results shown in Figure 7.24, fuel-switching for industries is very reliable with a small margin of differences between all sensitivity scenarios. Biomass fuel plays a dominant role in industrial process heating, with biomass fuel from 10.1 TWh per annum to 17.2 TWh per annum, whereas biomass end-use demand is from 7.7 TWh per annum to 13.5 TWh per annum which is around 85-90% of total decarbonised process heat. On the other hand, despite electrification being in range of 2.4 TWh per annum to 7.0 TWh of input electricity for process heat, the thermal load delivered would be 3.5 to 9.2 TWh per annum. The figure further shows that changes in biomass price have the greatest impact on fuel-switching compared with changes in electricity price and biomass availability. This effect is evident from the steep increase observed in the figure when biomass price is increased to 400%.

7.7 Limitations and Conclusions

In this chapter, a single deterministic optimisation energy system problem has been solved. This chapter presents key findings on the fuel-switching pathways for each industrial site within the studied boundary. These findings include capacity sizing decisions and operational flows at every time step for each heating technology across industrial sites. The highly spatial energy modelling framework also establishes energy commodity connections and network pathways between different system layers.

However, the high spatial resolution applied in the proposed energy-modelling framework introduces some limitations. The model contains 745,707 decision variables and 1,118,033 constraint equations, resulting in a large-scale optimisation problem. This scale makes it difficult to incorporate additional system components that may be important depending on the scope of analysis. The size of the problem not only increases computational solve times but also extends model construction and preprocessing time. As a result, solving a fully integrated national-scale optimisation within a single formulation would be computationally challenging.

It is acknowledged that announced industrial projects may reflect additional decision criteria not explicitly captured in the optimisation framework, such as corporate strategy, financing constraints, regulatory compliance requirements, and risk preferences. The optimisation model returns the optimised solution for the factory system, whereas the decarbonisation plans initiated by companies are developed according to their own strategic interests. Therefore, alignment with real-world projects is interpreted as consistency rather than exact prediction.

Moreover, the model has a limitation in that fuel-switching is conducted instantaneously in the projected year without considering the duration required for changeover. This limitation could be an important decision factor for companies that are concerned about potential interruptions to process operations during retrofitting activities, such as coal boiler conversions. Furthermore, site-level operational data is not publicly available for all facilities, limiting quantitative validation of exact capacity sizing or fuel consumption values. Validation is conducted at the technology selection and strategic pathway level.

This chapter treats electricity sales costs for industrial sites as exogenous assumptions, which represents a further limitation. While this approach enables overall system analysis of process heat fuel-switching decisions from the factory's perspective, it does not explicitly account for the cost of electricity system upgrades required to support electrification. In practice, investment in new generation, transmission, and distribution assets would increase system costs, and these costs would ultimately be reflected in electricity prices paid by consumers. Hence, the electricity cost assumptions applied here may underestimate long-term infrastructure-related cost impacts under high electrification pathways.

For this reason, this chapter focuses on solving the industrial process heat fuel-switching problem independently before introducing grid-specific constraints. Grid expansion upgrades, including transmission reinforcement, transformer capacity limits, distribution network impacts, and

electricity generation expansion, are deferred to the following chapter, where sequential deterministic optimisation is applied to quantify infrastructure-related costs.

In addition, conventional P-graph does not natively support a sequential optimisation approach in which the outcome of one optimisation is directly used as an input for a subsequent model. However, the development of the P-graph porting tool in this study enables systematic extraction of solution structures and connection edges from the first-stage model. These outputs are then transferred as parameters into the secondary grid-focused formulation. This approach reduces problem size and computational complexity while maintaining structural consistency across the system. In the next chapter, the sequential optimisation framework for assessing grid asset upgrades is presented.

Despite the deterministic multiperiod optimisation producing a good baseline assessment of fuel-switching pathways, the framework does not explicitly account for the electricity infrastructure expansion required for the increase in electrification. The increase in electricity demand for process heat and other non-industrial processes in the projected year would require additional generation capacity and changes in grid structural configurations. In order to address this limitation, the following chapter introduces a sequential optimisation framework that links industrial heat decarbonisation decisions with electricity system expansion.

Chapter 8: Sequential Optimisation

8.1 Introduction

This chapter extends the deterministic analysis by introducing a sequential optimisation framework to represent staged system development and the grid reinforcement requirements associated with large-scale industrial electrification. Sequential optimisation decomposes a complex decision problem into multiple interconnected stages, with each stage addressing a distinct set of decisions under progressively refined assumptions. In energy system modelling, this approach is well suited to problems involving infrastructure expansion, operational behaviour, and investment decisions that occur over different time scales and cannot be represented realistically within a single optimisation formulation without excessive complexity. The chapter establishes whether the Chapter 7 pathway remains near-optimal once electricity-system consequences are endogenised, which addresses the limitation from the factory-side model.

In the context of industrial decarbonisation, a key challenge arises from the interaction between process heat electrification and electricity system capacity. Large-scale electrification of industrial heat substantially increases electricity demand, which may place pressure on generation capacity and grid infrastructure. These pressures are likely to increase electricity supply costs as additional investment in generation, transmission, and distribution assets becomes necessary.

Many established energy planning tools, such as TIMES and MARKAL, typically relies on aggregated sectoral representations and scenario-based projections. While these tools are effective for exploring long-term energy transitions, they do not explicitly represent the infrastructure upgrades required to support large increases in electricity demand, nor do they directly capture how such upgrades influence electricity sales prices faced by industrial users.

This disconnect can lead to an underestimation of total system costs in pathways which relies heavily on electrification decarbonisation pathways. If the electricity prices are assumed to remain fixed at historical levels, the feedback between infrastructure expansion and electricity pricing is not accounted. However, in reality, although electricity sales prices are determined through complex and often confidential arrangements between generators, retailers, and industrial consumers, it is reasonable to assume that the costs of expanding electricity generation and grid infrastructure will ultimately be recovered through new consumer users.

To address this limitation, this chapter applies a two-stage sequential optimisation framework that explicitly links industrial process heat decarbonisation decisions with electricity system expansion costs. The first stage focuses on identifying least-cost fuel-switching and technology adoption pathways for industrial process heat under assumed electricity prices and fuel costs. These decisions represent the behaviour of industrial heat users responding to prevailing energy prices. The second stage then evaluates the electricity system implications of the resulting electrification outcomes, with particular emphasis on generation adequacy, grid capacity constraints, and the investment required to maintain system feasibility. This modelling also

includes the estimation of the percentage increase in electricity cost based on the upgrades needed.

Rather than solving a fully integrated expansion and operational problem within a single optimisation formulation, the framework adopts an iterative approach to maintain computational tractability. Solving both problems simultaneously would significantly increase model size and hinder convergence to a global optimum. Electricity prices used in the first stage are adjusted using a cost coefficient reflecting the additional infrastructure costs identified in the second stage. The updated electricity price is then fed back into the first stage, and the process is repeated until convergence is achieved between industrial electricity demand and the cost of supplying that demand.

8.2 Model Overview

The modelling approach adopted in this study is based on a two-stage optimisation concept. To the author's knowledge, this represents the first implementation of a P-graph to Pyomo sequential optimisation feedback architecture. The framework is designed to manage both computational complexity and the sequential nature of energy system transitions. The first stage, referred to as the factory model, focuses on the industrial process heat sector. The optimisation determines least-cost fuel-switching pathways for factory sites under a multiperiod operational structure. Technology switching and fuel dispatch decisions are chosen to minimise decarbonisation costs given assumed electricity prices and biomass fuel costs. This model corresponds to the optimisation framework presented in Chapter 7.

The second stage, which is the grid reinforcement stage, builds on the outcomes of the first by shifting attention from site-level technology switching to the electricity network implications of large-scale electrification. Instead of solving a fully integrated expansion problem in a single formulation, the framework treats grid reinforcement and infrastructure constraints as a downstream consequence of the process heat transition. In this stage, key decisions from the first optimisation, such as factory fuel-switching outcomes, biomass allocation, and methanol production levels, are fixed and are transformed from decision variables into parameters. By reducing the decision variables in the factory, biomass and methanol layers while expanding the second model layer to the grid representation in greater detail. Electricity prices are not optimised in the second stage. Instead, the model endogenously estimates infrastructure expansion costs, which are added to a fixed base sales price to reconstruct the electricity price used in the next first-stage iteration.

In this study, the electricity sales prices are assumed to reflect the aggregated cost of supplying electricity across all major demand sectors, including industrial process heat, non-process heat industrial demand, residential, commercial and transport. This assumption recognises that electricity infrastructure investments are driven by total system demand rather than by any single sector in isolation. As the studied year is 2037, growth in residential and commercial electricity demand, along with increased uptake of electric vehicles, is expected to contribute materially to overall electricity demand and associated infrastructure expansion. Therefore, it is appropriate to apply electricity price adjustments based on total system sales, rather than isolating industrial

electricity consumption alone when estimating the electricity sales price increases. This aggregated pricing assumption ensures consistency between electricity demand growth and infrastructure-driven price adjustments in the sequential optimisation framework.

Table 8.1 summarises the sectoral electricity demand, average electricity sales price, and resulting total electricity sales cost used as baseline inputs for the 2022 reference year and the 2037 projection year. These values form the basis for reconstructing total electricity sales cost in both stages of the sequential optimisation framework. The projected 2037 electricity demand reflects growth in commercial activity, residential consumption, and transport electrification, whereas existing industrial demand is kept constant except for the process heat adjustments determined in Stage 1.

Table 8.1: Summary of the sectoral electricity demand and average sales electricity price for base value (MBIE, 2024a).

Sector	Base Year 2022			Projected Year 2037		
	Demand (TWh yr^{-1})	Price (NZD MWh^{-1})	Cost (NZD yr^{-1})	Demand (TWh yr^{-1})	Price (NZD MWh^{-1})	Cost (NZD yr^{-1})
Existing Industrial $CC^{Ind,exist}$	13.5	19.55	1.71E+9	13.5	19.55	1.71E+9
Commercial & Agriculture CC^{Com}	11.8	21.16	2.49E+9	12.8	21.16	2.70E+9
Residential CC^{Res}	13.4	34.88	4.68E+9	14.8	34.88	5.17E+9
Transport $CC^{Transport}$	0.185	34.88	6.44E+7	2.32	34.88	8.08E+8
Total	38.9	-	9.87E+9	43.4	-	1.13E+10

A further assumption relates to grid reliability requirements. Due to the limited availability of detailed transmission planning data, the electricity network is modelled under an N security criterion rather than the more conservative N-1 security standard typically applied in transmission planning.

Under the N security assumption, the network is required to meet demand under normal operating conditions but is not explicitly designed to withstand the loss of a single major component without service disruption. As a result, the grid reinforcement and upgrade costs estimated in this study represent a lower bound on the true infrastructure investment required to support the projected level of electrification. In practice, enforcing an N-1 reliability standard would require additional redundancy in transmission lines, transformers, and generation capacity, leading to higher capital investment and electricity supply costs than those estimated here. This

assumption results in a conservative estimation of the upgrade costs for grid planning requirements.

The second stage explicitly incorporates POC, transmission interactions, and generation sites so that both supply availability and network capacity constraints can be assessed under increased electrification demand. The optimisation then concentrates on upgrades such as generation and potential reinforcement requirements needed so as to maintain feasibility and security. By separating industrial transition decisions from network and generation adequacy decisions, the staged approach remains computationally tractable while still capturing the structural dependence between process heat electrification, electricity supply, and grid infrastructure. Table 8.2 summarises the specifications and aspects that differentiate the scope between the two stages.

Table 8.2: Summary of decision scope and assumptions between the Stage 1 and Stage 2 optimisation framework.

Specifications	Stage 1	Stage 2
Purpose	Industrial process heat decarbonisation optimisation	Generation and grid expansion cost assessment
Electricity Price	Exogenous and iteratively updated	Reconstructed as base sales price plus endogenous infrastructure cost adjustment
Process Heat Industrial Technologies	Endogenously optimised	Removed from the Stage 2
Industrial electricity demand	Decision variable	Fixed input
Biomass allocation	Optimised across competing uses	Fixed input
Methanol Production	Optional and endogenous	Fixed input
Grid representation	Simplified	Explicit generation, transmission and distribution
Investment decisions	Industrial technologies with electricity and biomass cost	Generation and grid assets
Objective	Minimise industrial decarbonisation cost	Minimise grid expansion cost

8.2.1 Model Concept

At the backend of the modelling framework, the optimisation problem is constructed in Python using the P-graph library and subsequently translated into a Pyomo formulation for numerical solving. Once the first-stage optimisation is solved in Pyomo, the resulting solution is extracted and used to restructure the model for the second stage.

Within this transition, the biomass layer is simplified. Biomass-related solution variables are removed, except for those nodes that define the thermal electricity generation structure. This ensures that the physical feasibility of biomass-based electricity generation is preserved while preventing biomass costs and flows from being re-optimised again in the second stage. In parallel, the industrial process heat layer is decomposed and aggregated (Figure 8.1). This simplification is intentional, as the objective of the second stage is to assess grid upgrade requirements relative to the base electricity price rather than to revisit industrial technology choices.

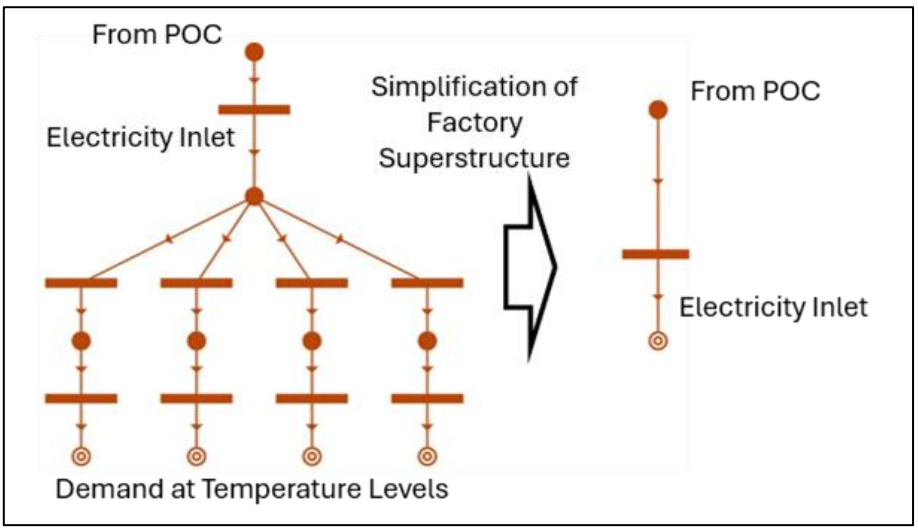


Figure 8.1: Illustrative example showing how the factory layer is simplified between stages in the multistage optimisation approach.

On the methanol side, biomass flows are removed from the second stage formulation, while the methanol production level obtained in the first stage is retained as a fixed parameter. This methanol product volume acts as a proxy for competing electricity demand, ensuring that electricity generation sites face consistent resource competition without reintroducing biomass allocation decisions.

Additional modelling enhancements are included in the second stage to improve computational performance and structural realism. These include the predetermined distribution cost functions for industrial electricity line upgrades, as well as informed initialisation of the number of identical grid reinforcement units included in the superstructure. This approach reduces problem size and improves solvability for the transmission layer.

In the second-stage optimisation, electricity generation expansion and power flows across generation, transmission, and distribution layers are re-optimised to identify the least-cost feasible grid configuration subject to network and capacity constraints.

Electricity generation costs associated with green methanol production are excluded from the electricity price and grid upgrade cost calculation for industrial and non-industrial demand. This exclusion is because the purpose of this stage is to estimate an updated electricity sales price for industry, which is subsequently passed back to the first stage and directly influences fuel-switching decisions.

The electricity generation layer is expanded to include additional network components. These include supply transformers, sub-transmission lines for new generation sites, demand transformers, distribution lines, and transmission network lines. Each of these components carries associated capital costs, which collectively increase the overall electricity supply cost. In contrast to the first stage, costs that were previously omitted are explicitly included in the second stage. These include new electricity generation costs, biomass resource costs, and biomass transport costs. Methanol product prices are excluded from the calculation of new generation costs in this stage through the use of a product node that cancels the generation costs while preserving the overall model structure. Figure 8.2 illustrates the single time-step energy system with grid reinforcement model.

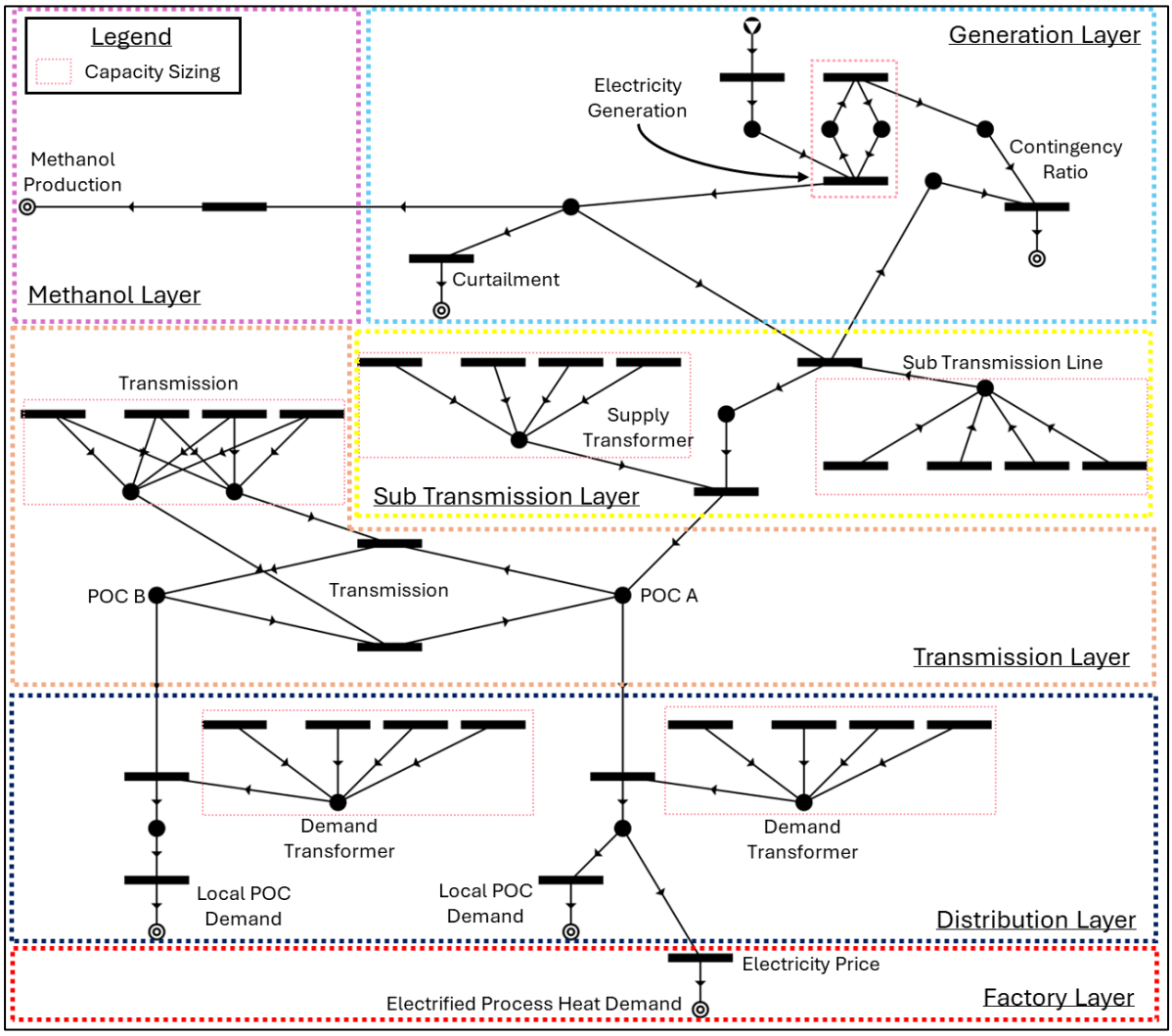


Figure 8.2: Illustration of a single time-step energy system with grid reinforcement, consisting of one generation unit, two POCs, one methanol plant, and one industrial factory.

8.3 Sequential Optimisation Modelling Framework

The sequential optimisation begins by setting up the root-finding method and solver environment, including configurations such as tolerance, maximum iterations, and the initial $\alpha^{e\ adjusted}$. The first step is to run the factory model, while the second step is to run the grid reinforcement model using the results from the factory model. The electricity sales costs are then computed, and if the solution has converged, the loop ends. If it has not converged, the first stage is rerun with a new $\alpha^{e\ adjusted}$. Once convergence is achieved, the final value of $\alpha^{e\ adjusted}$ is returned together with the root-finding results and the optimisation results for both model stages. The solution is then saved and the operation is terminated. Figure 8.3 outlines the architecture of the two-stage optimisation model.

The reduction and simplification to two models are intentional to decrease the problem size and shift the modelling emphasis towards grid modelling. In this framework, the detailed factory superstructure is aggregated into a demand node, while the second stage assesses grid expansion by introducing distribution line capacity sizing. This approach demonstrates the capability of the energy modelling framework presented in the thesis, which enables one large problem to be decomposed into two smaller and more manageable problems.

The two stages are then linked through an iterative cost updating procedure. After each second-stage solve, the estimated grid upgrade cost is combined with the base electricity sales price. If this combined cost differs from the electricity price assumed in the first stage, the electricity price coefficient $\alpha^{e\ adjusted}$ is updated, and the first stage optimisation is rerun. Figure 8.3 outlines the architecture of the two-stage optimisation model. Convergence is defined when the reconstructed total electricity sales cost in Stage 2 is approximately equal to the coefficient adjusted total electricity sales $\alpha^{e\ adjusted}$ which was assumed in Stage 1.

The sequential optimisation framework requires careful treatment of the component structure, particularly the role of representative nodes within the model. In the first stage of the optimisation, electricity costs are estimated by setting the price across all POCs, reflecting the fact that each POC exhibits a different electricity price structure within the New Zealand electricity market, as it is influenced by the wholesale electricity price. The primary objective function of this first stage is to minimise fuel-switching costs for process heat industries within the system model.

To support these calculations, the electricity generation component is fully represented in the first stage of the model. This includes capacity installation decisions, site location selection, and the availability of biomass resources for thermal electricity generation. Incorporating these elements at this stage establishes a consistent initial structure that defines how biomass resources can be connected to thermal generation, while enforcing resource limits across competing uses such as industrial heating and green methanol production. The underlying operational costs of electricity generation are not explicitly observed as electricity prices, since these costs are embedded within the electricity sales prices determined by electricity companies.

However, electricity sales prices are not static. In addition to differing across POCs, they are expected to increase if significant upgrades or new electricity generation infrastructure are

required to meet rising demand from industrial electrification, increasing electric vehicle uptake, and projected population growth by 2037. The multistage optimisation framework is therefore used to capture both the adoption of net zero industrial heating technologies and the resulting adjustments in electricity sales prices driven by infrastructure expansion and new investment costs.

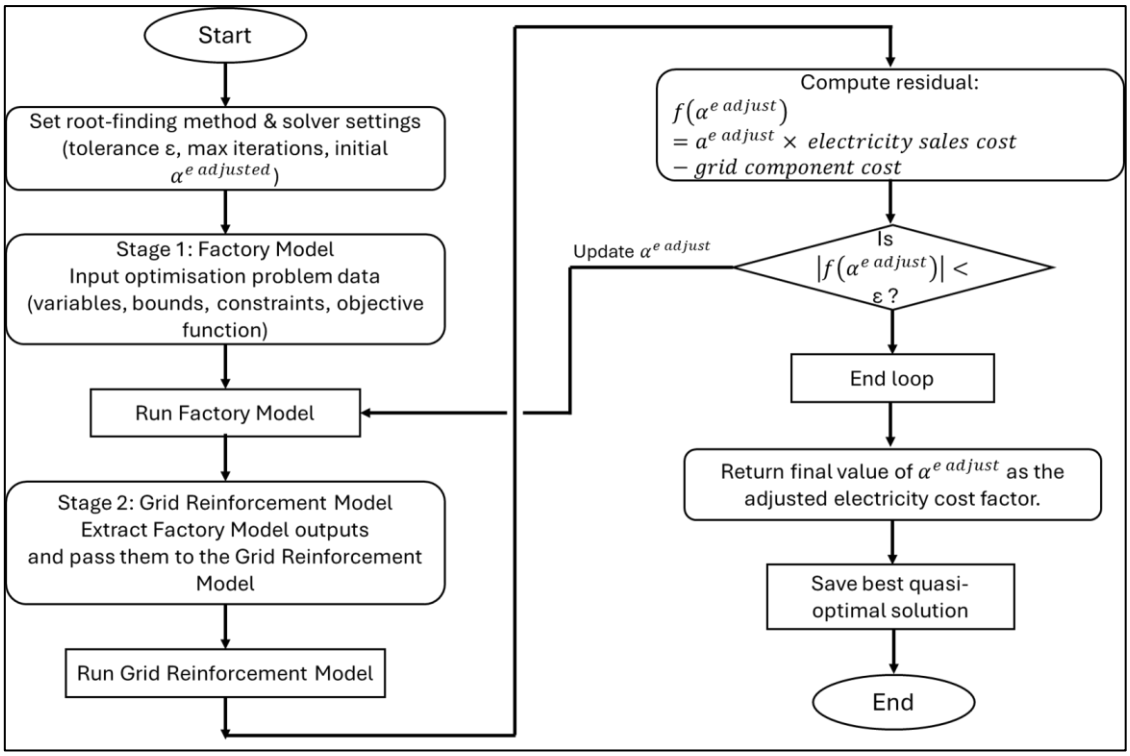


Figure 8.3: Model architecture of the two-stage optimisation model.

In order to automate the search for the electricity price coefficient $\alpha^{e\ adjusted}$, the problem is formulated as a root-finding problem. The root-finding algorithm allows relevant cost components in the first stage to be matched with those in the second stage until both values converge. In this case, the technique used to resolve convergence can be implemented using the SciPy library. The mechanism of the sequential structure is as follows: the framework has an outer-loop problem in which the electricity price coefficient $\alpha^{e\ adjusted}$ is iteratively updated until the stages converge. As a result, a scalar nonlinear fixed-point problem is defined over the electricity price adjustment parameter.

To solve the problem efficiently, Brent's method, which is available in the SciPy library, is employed. It is chosen because it combines the fast convergence properties of the secant method with the robustness of the bisection method, which guarantees convergence if the root is properly bounded. This makes it suitable for this energy modelling framework application, where the objective landscape is piecewise due to the nature of the MILP formulation used in both stages.

Although each stage is comprised of discrete decision variables and linear constraints, the mapping between the electricity price adjustment coefficient and system cost is continuous. Thus, a derivative-free root-finding algorithm such as Brent's method is selected.

8.4 Model Formulation

Sets

While maintaining the sets and indices from the previous chapter, additional sets are introduced to represent the electricity network infrastructure in their physical asset types. This formulation allows transmission constraints, upgrade decisions and capacity ratings to be applied consistently across the grid modelling layer. As grid infrastructure expansion is modelled with the help of discrete asset options rather than continuous capacity scaling, the auxiliary grid assets are described by sets where the elements represent the predefined equipment sized. This approach reflects the standardised nature of real-world electricity infrastructure. The sets and indices which are introduced in the section includes, transformers S^{tr}, transmission lines V, and sub-transmission along with distribution lines R^d.

Transformer expansion is represented through a discrete set of standard capacity options, which apply to both supply-side and demand-side transformers, which are defined as:

$$s \in S^{tr} \qquad (8.1)$$

where each element s corresponds to the transformer specification and capacity within the supply transformer and demand transformer system. The transformer index is defined according to Appendix D, ordered in ascending capacity rating.

The asset type sets provide the structural foundation for the heterogenous grid infrastructure. The transmission lines formulation enables the assignment of cost parameters with capacity constraints and reinforcement options depending on the characteristics of the line. The transmission lines are distinguished as the following:

$$v \in V \qquad (8.2)$$

where v represents the specifications of the lines based on voltage level, circuit configuration and capacity limits.

Similarly, distribution and sub-transmission line types are included to represent the connections between generation sites and POCs as well as between POCs and downstream distribution demand nodes. These line capacity classes are denoted by:

$$r^d \in R^d \qquad (8.3)$$

where r corresponds to the line type associated with the power rating.

In short, these sets of grid assets provide the suitable foundation for representing heterogenous grid infrastructure technologies embedded within the second stage optimisation, while maintaining the consistency with the discrete sizing approach adopted in the P-graph framework.

Objective Function

In order to be consistent with the two-stage sequential optimisation framework described in Section 8.2 and the Figure 8.3 which describes the model architecture of the energy modelling

framework, separate objective functions are defined for the industrial process heat optimisation and the electricity system upgrades optimisation. Each objective reflects the decision-making context of its respective stage, while the interaction between the two is managed through an iterative updating mechanism.

First Stage Objective

In the first stage, the objective function represents the cost minimisation problem encountered by the industrial process heat users. The optimisation seeks to minimise total industrial decarbonisation cost, which includes process heat technologies, fuel supply, transport and operating costs under the assumed electricity price. Electricity system expansion structure and costs do not appear explicitly in this stage. Instead, Stage 1 uses the base sales price plus an iterative multiplier where a new electricity cost coefficient, denoted as $\alpha^{e\ adjusted}$ that approximates the upgrade cost pass-through estimated in Stage 2. It is represented as a grid cost proxy where a new electricity cost coefficient, denoted as $\alpha^{e\ adjusted}$ which would be a proxy of what the adjusted electricity cost is, considering the upgrades.

This section extends the process heat optimisation by introducing a second-stage grid reinforcement model. The objective function is separated into two sequential optimisation stages. The first-stage objective is formulated in Equation (8.4) as:

$$\min CC^{bio(1)} + CC^{MeOH(1)} + CC^{e,st(1)} + CC^{bio,fac,trans(1)} + CC^{bio,MeOH,trans(1)} + CC^{ht(1)} + \alpha^{e\ adjusted} CC^{salescost,PH(1)} - P^{MeOH(1)} \quad (8.4)$$

Second Stage Objective

In the second stage, the objective function outlines the electricity system response required to supply the electricity demand indicated by the first stage outcomes. Industrial technology choices and electricity demand are treated as fixed inputs, and the optimisation focuses on identifying the least-cost combination of electricity generation expansion and grid reinforcement investments required to maintain system feasibility.

The cost function for the second stage optimisation includes upgrades on:

- new grid-connected electricity generations $CC^{gen(2)}$,
- supply transformers $CC^{tr,\mathrm{sup}(2)}$,
- demand transformers $CC^{tr,dem(2)}$,
- transmission lines $CC^{trans,line(2)}$,
- sub-transmission lines $CC^{subtrans(2)}$ and
- distribution lines $CC^{dist(2)}$.

The second stage involves the grid reinforcement modelling which is expressed as:

$$\min CC^{grid,upgrade(2)} \quad (8.5)$$

The decomposition of the total grid upgrade cost is comprised of transmission, transformer, distribution, and generation investment components.

$$CC^{grid,upgrade(2)} = CC^{trans,line(2)} + CC^{tr,\sup(2)} + CC^{tr,dem(2)} + CC^{subtrans(2)} + CC^{dist(2)} + CC^{gen(2)} \quad (8.6)$$

Iteration between Stages

The two objective functions are linked through an iterative consistency condition that is a term within the electricity sales cost assumed in Stage 1 with the infrastructure cost estimated in Stage 2. After each second stage optimisation, the estimated grid reinforcement cost is compared with the grid cost proxy assumed in the first stage. The grid cost proxy is then updated to reflect the infrastructure investment required to supply the resulting electricity demand.

The total electricity sales cost in Stage 2 is the sum of grid upgrade costs and the base electricity sales cost shown in Equation (8.7). Equation (8.8) expresses the total electricity sales cost based on different sectors of electricity demand and tariff levels. Equation (8.9) describes the breakdown of the base electricity sales cost originating from residential, commercial, transport, and existing industrial demand sectors.

$$CC^{e\ sales\ cost(2)} = CC^{grid,upgrade(2)} + CC^{base} \quad (8.7)$$

$$CC^{e\ sales\ cost(2)} = CC^{grid,upgrade(2)} + \sum_{sector \in SECTOR} C_{sector} D_{sector}^{base} \quad (8.8)$$

$$CC^{base} = CC^{Res,base} + CC^{Com,base} + CC^{Transport,base} + CC^{Ind,exist,base} + CC^{unallocated} \quad (8.9)$$

In Stage 1, the new electricity cost coefficient $\alpha^{e\ adjusted}$ is applied to estimate the total electricity cost.

$$CC^{e\ sales\ cost(1)} = \alpha^{e\ adjusted} \left(CC^{gridcost,PH(1)} + \sum_{sector \in SECTOR} C_{sector} D_{sector}^{2037} \right) \quad (8.10)$$

where $\alpha^{e\ adjusted}$ is initialised as 1.0 in the first iteration, and it is updated iteratively until convergence is achieved, such that the grid cost proxy used in the first stage matches the electricity system reinforcement cost identified in the second stage. The increase in electricity demand in the non-industrial sectors also contributed to the grid upgrades, hence, assuming the price increment is done uniformly for each sector within the country, the $\alpha^{e\ adjusted}$ would be such that:

$$\alpha^{e\ adjusted} = \frac{CC^{grid,upgrade(2)} + \sum_{sector \in SECTOR} C_{sector} D_{sector}^{base}}{CC^{gridcost,PH(1)} + \sum_{sector \in SECTOR} C_{sector} D_{sector}^{2037}} \quad (8.11)$$

This iterative process is done using Brent's method until convergence is achieved, such that the grid cost proxy used in the first stage matches the electricity system reinforcement cost identified

in the second stage. The absolute tolerance and relative tolerance are set as 0.001 and 1X10^{-6} respectively to increase the accuracy of the $\alpha^{e\ adjusted}$ value.

$$CC^{e\ sales\ cost(1)} \approx CC^{e\ sales\ cost(2)} \quad (8.12)$$

This condition ensures that the electricity sales price assumed by the industrial users for process heat decarbonisation as well as electricity demand growth in non-industrial sectors in the first stage is consistent with the required grid infrastructure improvements to support the level of electrification suggested in the second stage.

8.5 Output and Reporting

8.5.1 Model Size and Computational Performance

The model size for the second stage is discussed here, as the factory model would have the same problem size and solving time as in the previous chapter, given that the scope of the study remains the same. Table 8.3 shows the size of the model in Stage 2, which represents the grid reinforcement model. Within the P-graph framework, there are a total of 32,909 material nodes and 28,609 operating unit nodes. This indicates that the model size is smaller than that of the factory model, due to the simplification of the biomass connection network as well as the onsite process heat configurations.

In summary, there are a total of 155,970 directed edges connecting each node in the model. Once the P-graph model is ported into Pyomo, the problem size can be quantified in terms of mathematical formulation. The model contains 359,179 decision variables, 508,979 constraints, and 120,826 binary variables. Each iteration takes approximately 171 seconds to solve, with a total elapsed time of about 280 seconds for defining, building, and exporting the problem.

Table 8.3: Summary of optimisation model size and computational performance for the grid reinforcement model.

Metrics	Values
Decision Variables	359,179
Constraints	508,979
Binary Variables	120,826
Continuous Variables	238,353
Time for defining the problem (s)	17
Time for building the problem (s)	92
Time for solving the problem (s)	171
Total elapse time (s)	280

8.5.2 Convergence Root Finding $\alpha^{e\ adjusted}$

Convergence on the solution root, the solution root or $\alpha^{e\ adjusted}$ in this scenario is 0.9065. The number of iterations is 4 with the absolute tolerance of 0.001 and relative tolerance of $1X10^{-6}$ of before converging into a solution. As the $\alpha^{e\ adjusted}$ is determined to be 0.9065 such that electricity costs in both the factory model and grid model converge. This implies that the electricity proportional cost change is -9.35% after the upgrades, assuming an N demand configuration and assuming the electricity price drops uniformly across all sectors. It indicates that the deployment of renewable technologies in the future would reduce the electricity price (Boston Consulting Group, 2025). The model is aligned with the report, which concludes that the electricity price would decrease in the long run, where the industrial consumer bills and residential consumer bills could decrease with increasing renewables and a higher demand response.

Further checks have been conducted on the Class 5 cost estimates, so there is a large uncertainty in the electricity infrastructural upgrades, so a wide accuracy range of -50% to +100% based on the cost is calculated to study the effect of electricity price change on the uptake of electrification. Hence, the sequential optimisation is conducted for both instances with different price ranges. Based on Table 8.4, it is estimated that the electricity cost would be determined and would be consistent where the main driver for fuel switching in decarbonisation would be biomass resources.

Table 8.4: Uncertainty analysis based on grid infrastructure cost.

Cost Range	Iterations	Electricity Price Change	Electricity Required (TWh yr^{-1})	Electricity for End-Use (TWh yr^{-1})
-50%	10	-12.6%	2.84	4.50
Base Scenario	8	-9.1%	2.81	4.45
+50	7	-5.5%	2.78	4.39
+100	9	-2.0%	2.76	4.34

8.6 Output Structure and Discussion

In this section, key results which are presented in this chapter include primary emphasis on the second stage model, which is the grid reinforcement model. The discussion focuses on the outcomes regarding the generation expansion, sub-transmission lines, transmission reinforcement, POC upgrades, and distribution network investments. These components collectively represent the implications not only of the electrification of industrial process heat but also the increase in electric vehicle uptake, residential and commercial in the projected year.

In addition, the selected results from Stage 1 are briefly revisited to compare the price different adjusted by $\alpha^{e\ adjusted}$ to account for infrastructural upgrades costs. This comparison highlights how the adjusted electricity price coefficient $\alpha^{e\ adjusted}$ would ultimately influence the electricity price outcomes as well as the scale of grid infrastructure upgrades. By examining the differences

between the two stages, the discussion clarifies how cost feedback mechanisms alter fuel-switching decisions and system-wide investment requirements.

8.6.1 Electricity Generation Layer

Based on the results in Figure 8.4, the largest share of generation upgrades consists of wind, geothermal, and hydro, with a relatively small contribution from solar. The model suggests installation of approximately 5.9 GW of new generation capacity in total. The addition of new generations caters for the increase in electrification uptake and also offsets the reduced utilisation of existing thermal power plants, particularly natural gas and coal generators. The model structure prioritises the direct use of biomass for industrial process heat rather than increasing electrification to displace thermal generation indirectly. As a result, biomass is allocated efficiently to heat applications at industrial sites, reducing pressure on electricity generation expansion. Figure 8.5 displays the spread of existing and proposed generated sites in New Zealand.

The model incorporates all proposed geothermal plants in the Waikato and Bay of Plenty regions, contributing approximately 1022 MW of installed capacity. In Figure 8.4, geothermal resources are geographically exclusive in certain regions. Therefore, there is no available geothermal expansion project observed in other parts of the country. Hydroelectric upgrades contribute approximately 110 MW. The relatively modest hydro expansion indicates that New Zealand has already developed a significant portion of its economically viable hydro capacity. The hydroelectricity expansion in the projected year is rather limited in terms of the expansion potential.

Wind generation accounts for a large share of the new capacity, with a total of 4719 MW installed across the country. Onshore wind is deployed due to the suitability of the region and due to the high wind conditions to replace thermal generation which is uneconomic under the assumptions.

Solar generation plays a comparatively smaller role based on the optimised results. While northern regions exhibit higher capacity factors, the model includes limited documented large-scale solar projects, and much of the existing development pipeline consists of local or distributed solar installations. At a monthly temporal resolution, distributed solar contributes less effectively to peak demand management and system stability. As a result, the optimisation structurally favours more stable and dispatchable generation technologies.

In the model, local or behind-the-meter solar is not treated as a simple reduction in local demand. Instead, it is represented as utility-scale solar connected to the grid at defined generation nodes. This modelling choice allows solar generation to be explicitly included within the generation layer and connected through POC nodes, thereby participating in system-wide electricity balance and infrastructure loading.

By modelling local solar as grid-connected generation rather than behind-the-meter demand reduction, the electricity flow is contributed to transformer loading, POC balance and distribution

connections. Electricity generated from solar would pass through the network structures and it influences the necessary upgrades for substations, sub transmission and distribution layers.

This abstraction supports structural consistency within the modelling framework, because electricity demand is aggregated at the POC level, making it difficult to explicitly account for behind-the-meter generation. Representing distributed solar as grid-connected generation therefore preserves its system-level impact on capacity expansion decisions without introducing additional structural complexity. Figure 8.6 outlines the grid-connected new generation sites to the nearest POC.

Under the modelling assumptions, the results indicate that the renewable generation expansion and the displacement of thermal generation outweigh the additional grid reinforcement. The primary generation technologies selected exhibit low marginal operating costs compared to the fossil-fuel-based thermal plants. This is because under the modelling parameters, a cost is incurred for retrofitting conventional thermal turbine units to biomass-powered ones, whose biomass resources, which are represented, would incur a large amount of expense in the long run. As a comparison, the cost of wholesale imported sub-bituminous Indonesian coal and natural gas is 7.5 NZD GJ^{-1} and 9.54 NZD GJ^{-1} in 2022, whereas k-log wood costs 16.4 NZD GJ^{-1} excluding transport costs, and it is a competing resource for on-site industrial heating (MBIE, 2024a). Moreover, fuel costs are usually susceptible to volatility, where sensitivity analysis of a high biomass cost or low availability due to increasing demand would be a good risk assessment.

The model suggests expanding long-term renewable generation technologies in the optimised system, under the projected 2037 technology cost assumptions. As a result, the levelised cost of new renewable generation is sufficiently competitive that the fuel cost savings from displaced thermal generation more than compensate for the incremental capital costs required for both generation expansion and network reinforcement. This is supported by the results, where there is a reduction in average electricity sales cost with the modelling assumptions being deployed in the case study. This demonstrates that the relationship between electrification and electricity price is not proportional with one another, as it is dependent on the deployment of capital-intensive but marginal-cost generations with fuel-intensive thermal technologies.

Given that the primary research question focuses on identifying optimal fuel-switching technologies for decarbonisable industrial process heat, detailed modelling of decentralised generation is beyond the scope of this thesis. Hence, a more granular representation of distributed solar and feeder-level effects could be incorporated in future work as an extension of this modelling framework.

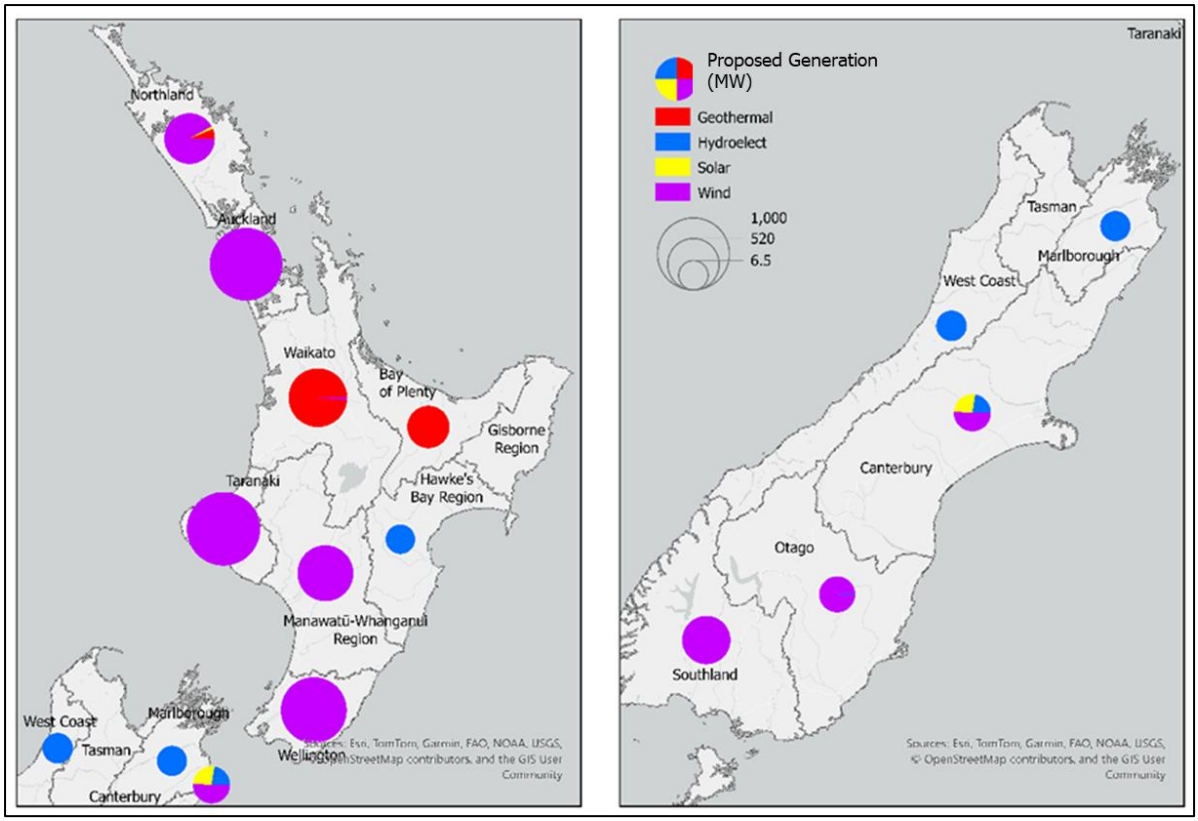


Figure 8.4: Installed new generation capacity by technology and region.

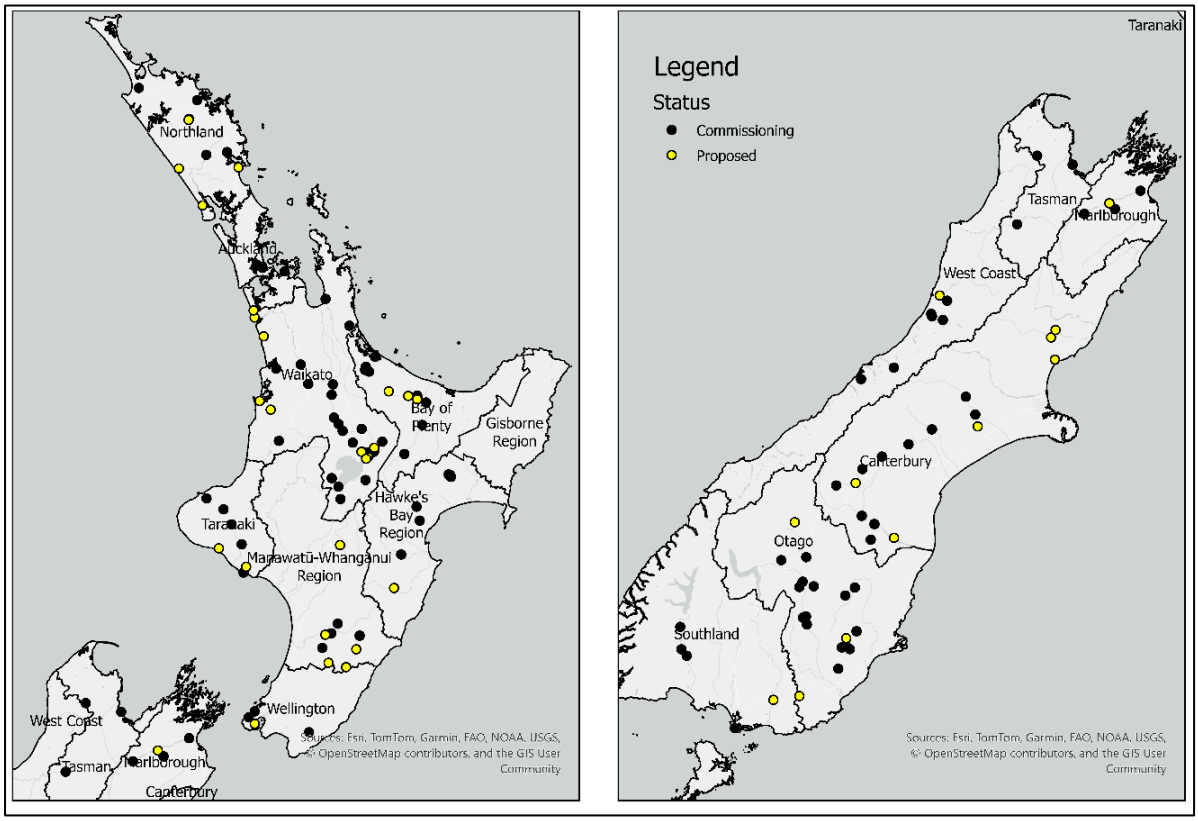


Figure 8.5: The generation portfolio including existing and proposed assets implemented in the model solution.

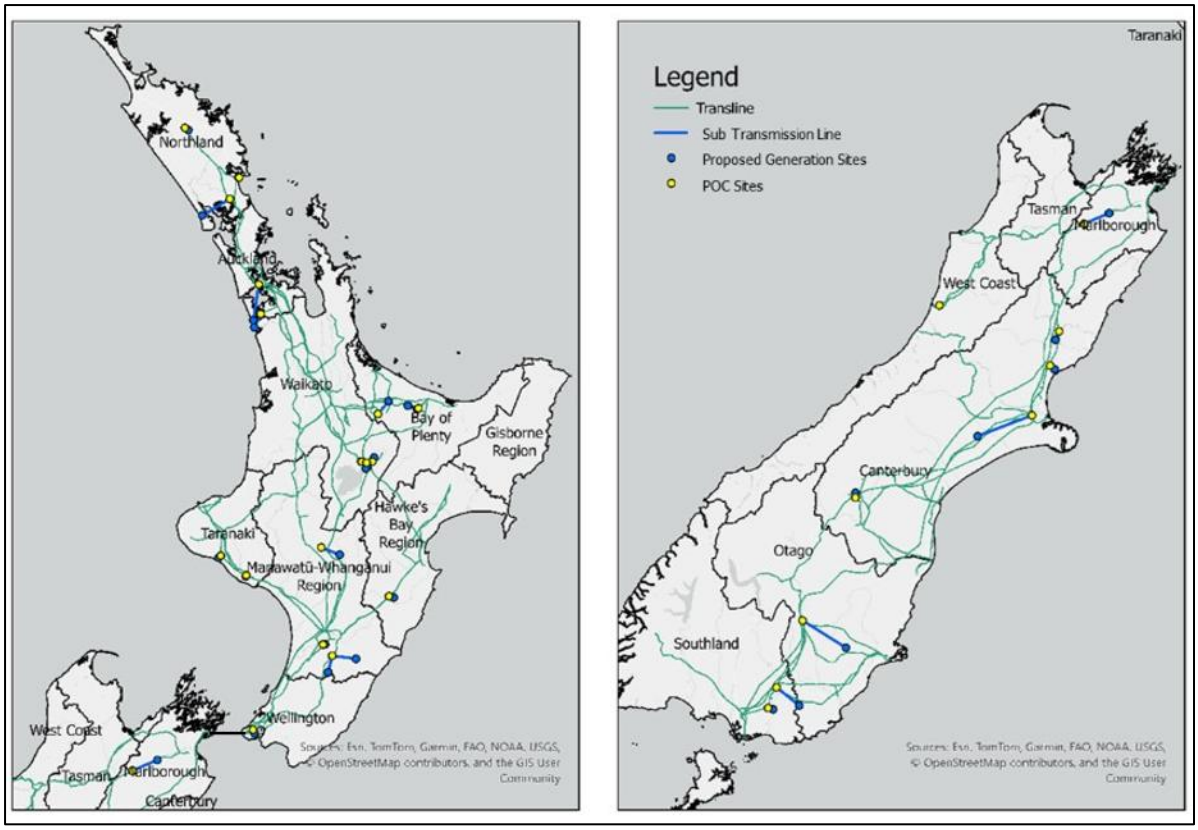


Figure 8.6: Sub transmission network expansion from newly installed generation to POC nodes.

8.6.2 Transmission Layer

For grid transmission modelling, only selected key transmission grids require upgrades. The reason is due to grid limitations related to electrification debottlenecking within the network. Figure 8.7 presents the spatial distribution of transmission network upgrades across the country. As the transmission grid is also unidirectional at every time block, the situation is also driven by increases in non-industrial electricity demand.

Several projects have been announced for the transmission grid, such as the Arapuni–Hamilton-B line and the Bombay–Meremere-A line. From the RETA report prepared by Ergo Consulting (2024), the mitigation plan for the bottlenecking in the ARI-HAM-B line is to create a permanent system division as well as busing the Arapuni–Bombay circuit to disconnect the Hamilton–Bombay connection. Hence, the proposal would inherently reduce the load within the transmission line without requiring technical upgrades to the Arapuni–Hamilton-B line.

On the other hand, assessments have been made on the BOB–MER-A circuit as stated in the Transpower Planning Report (Transpower New Zealand, 2022b). Therefore, refurbishment projects would be carried out so that the line would no longer be a critical component by 2030.

Although the grid reinforcement model does not identify any upgrades for the Arahura–Otira-A line and the Whanganui–Stratford-A line, projects on adjoining lines have been observed,

particularly on the Coleridge–Castle Hill–Arthur's Pass–Otira–1 and Bunnythorpe–Marton–Whanganui lines. The detailed specifications regarding these upgrades are outlined in Table 8.5. However, there are no news or reports regarding upgrades for the Gore–Halfway Bush-A line. This could be because the projected year in the model is 2037, meaning it may not yet be considered a critical component within the next five years.

As a summary for this section, the grid reinforcement model provides insight into transmission line upgrades by enforcing electricity flows within grid constraints. Nevertheless, practical planning decisions still require human judgement, such as constructing new POC stations, establishing new transmission lines, or repairing poles and structures that have exceeded their lifespan, which are all not accounted for in the simplified model.

Table 8.5: Anticipated transmission network upgrades relevant to modelling results.

Lines	Upgrades Announced	Source
ARI-HAM-B	To establish a permanent system division at Ongarue on the 110 kV Hangatiki-Rangitoto Hills Ongarue circuit, thereby reducing the load on the Arapuni-Hamilton circuits. To eradicate the post-contingency overloads within the TPR forecast period, the Arapuni-Bombay circuit will be bussed at Hamilton, with the Hamilton-Bombay portion being disconnected. This will increase the capacity between Arapuni and Hamilton.	Ergo Consulting (2024)
AHA-OTI-A	The adjoining line Coleridge–Castle Hill–Arthur's Pass–Otira–1 may cause the remaining circuit to overload at the end of the forecast period.	Transpower (2023a)
BOB-MER-A	Transpower is undertaking a project in the investigation and delivery phase to enhance the power supply to the Bombay station. The line will no longer be a critical component by 2030 at the latest.	Transpower (2022b)
GOR-HWB-A	N/A	
WGN-SFD-A	Based on the Transpower possible 2038 configuration plan, upgrades of the 110 kV Bunnythorpe-Marton-Whanganui lines are needed, which is are adjoining lines to the WGN-SFD-A circuit.	Ergo Consulting (2025)

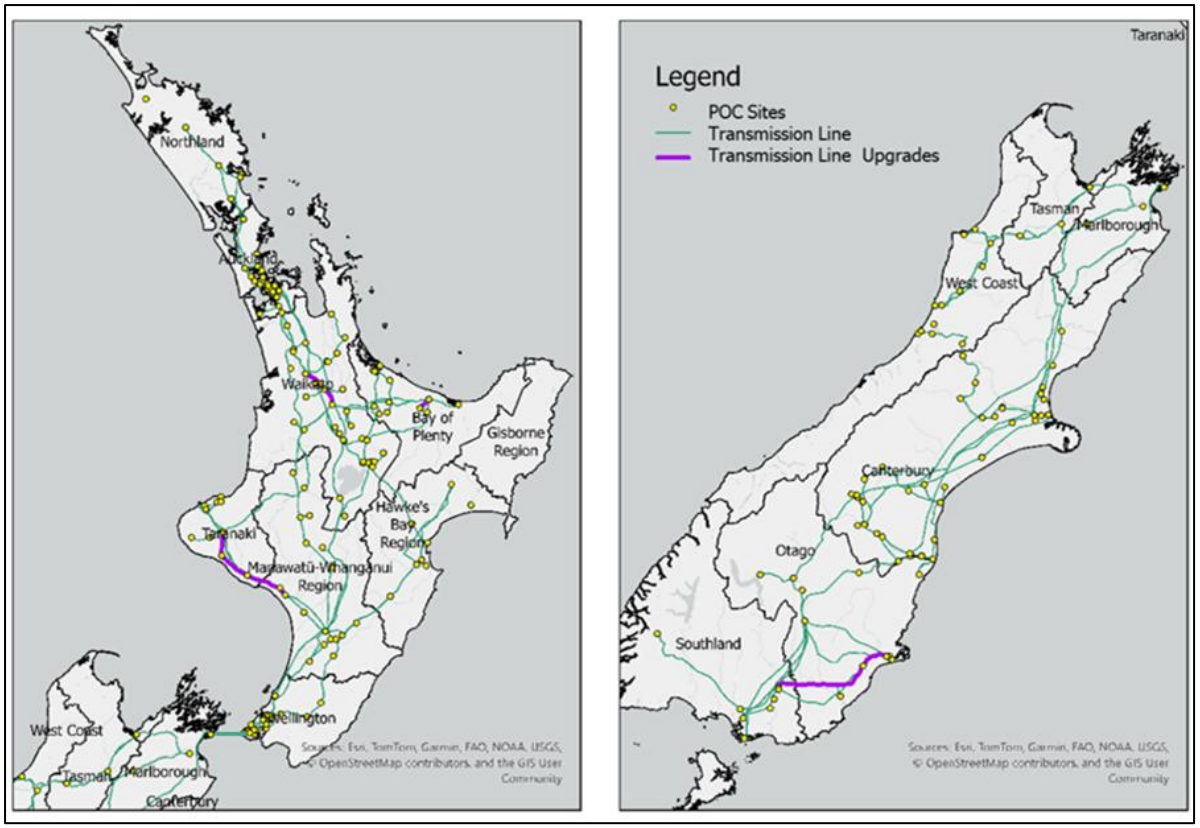


Figure 8.7: Newly reinforced and existing transmission network assets.

8.6.3 Distribution Layer

The distribution layer accounts for approximately 2.1% of the total capital cost of grid upgrades in the sequential optimisation results. Although this share is relatively small compared to generation and transmission investments, distribution reinforcement remains necessary in locations where electrification of industrial process heat increases local peak demand. In the model, distribution upgrades are triggered when the required electricity flow from a POC to an industrial site exceeds the available connection capacity. The reinforcement requirement is determined based on the maximum MVA capacity required across different time periods rather than annual energy volume. This consideration reflects the importance of peak loading in distribution asset sizing. The cost of each upgrade is therefore driven by the incremental MVA capacity required.

It is assumed that distribution capacity is expanded in discrete integer increments and that upgrade costs depend on both the required additional capacity and the distance between the industrial site and the associated POC. This formulation captures the spatial component of feeder extensions and conductor reinforcement within the distribution layer. There are 298 distribution line connections which are upgraded or newly built in the converged solution as displayed in Figure 8.8. Those upgrades contribute to accommodating additional electrification for process heat purposes at industrial sites. These upgrades occur primarily in regions where electrification

pathways are selected and where existing connection capacity is insufficient to support the increased load.

The relatively small capital share of the distribution layer indicates that, under the model assumptions, reinforcement costs are more significant at higher voltage lines such as the transmission and sub transmission lines. Moreover, the electrified industrial sites selected are generally located at short distribution line distances from the POC, leading to a low-cost contribution to the total reinforcement upgrade costs. This outcome suggests that total system costs are driven predominantly by bulk power transfer and generation expansion associated with system-wide electrification, rather than by local distribution constraints within the aggregated representation adopted in this framework.

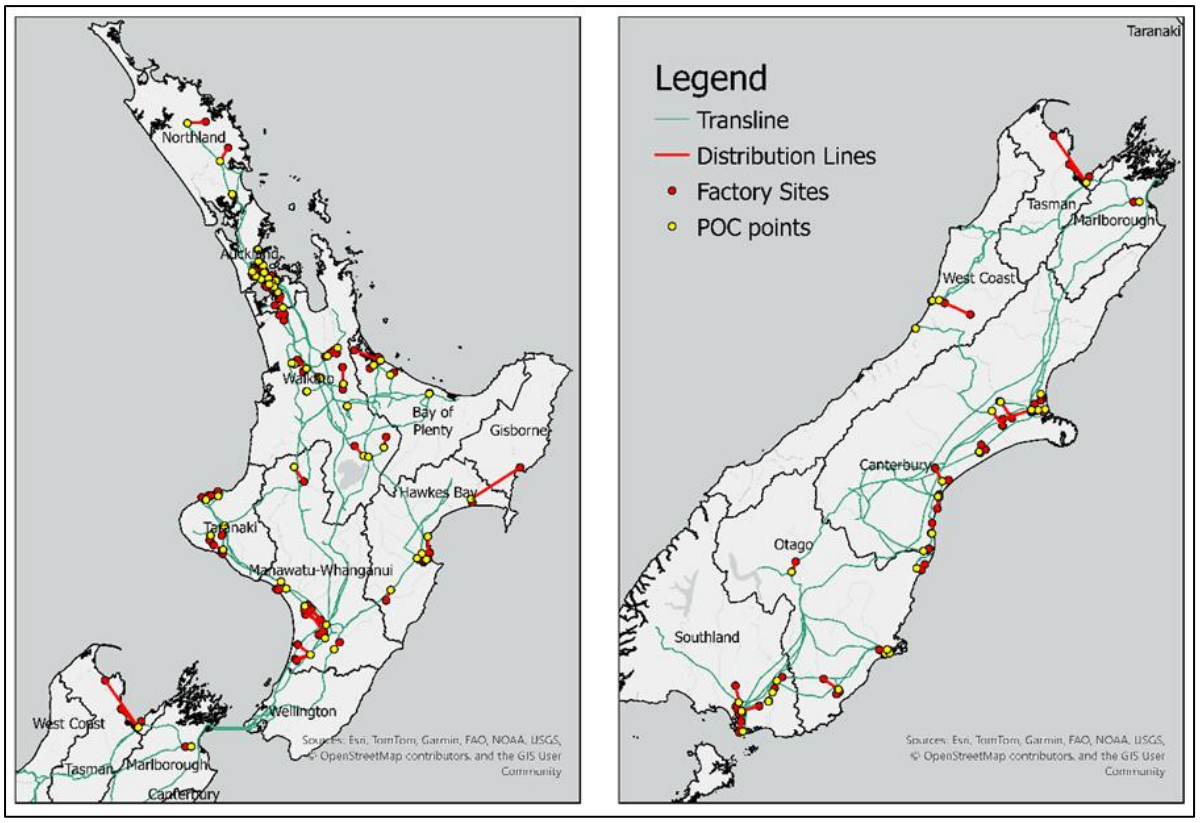


Figure 8.8: Distribution lines generated from the electrified process heat in industries to the POC nodes.

8.6.4 Summary of Grid Reinforcement Layer

The overall capital cost breakdown of infrastructure upgrades is obtained from the converged sequential optimisation framework. From Figure 8.9, the result from the model constitutes that the largest share of the reinforcement cost is generation, which is 82.4% of total annualised capital expenditure. This reflects the significant investment required to replace thermal generation and support additional electrification demand across the system.

Sub transmission upgrades account for 6.9% of the annualised total capital cost, while transmission reinforcement contributed 5.1%. These investments are required to transfer

electricity from newly installed generation sites to load centres and to maintain system adequacy under increased electrification scenarios. For the POC layer, supply transformer upgrades account for 2.5% of total capital costs whereas the demand transformer upgrades account for 1.0% of total capital cost, which represents the site-level electrical upgrades. Distribution network lines represent 2.1% of the total upgrade costs, which reflect the electricity grid connecting POC to the industrial sites.

The cost composition has several implications. Although grid reinforcement is required to accommodate additional electricity flow within fixed capacity limits, the dominant cost driver is the total investment incurred for generation capacity expansion. This problem arises because the model does not consider retrofitting fossil-based turbine generators to operate on solid biomass fuel. Instead, these units are replaced with newly installed renewable technologies that exhibit minimal variable operating costs. As a result, the new generation configuration reduces the overall system cost.

Since the objective function of the grid model is formulated to minimise total grid reinforcement and generation costs subject to resource and flow constraints, the optimisation returns the least-cost configuration that satisfies demand within the defined superstructure. The results indicate that installing new generation capacity and modifying the grid configuration is more cost-effective than relying on existing thermal generation assets and legacy network structures that are not aligned with rising electricity demand.

Overall, the capital cost distribution dominates system investment requirements, while the network-level reinforcement in the transmission and distribution levels accounts for a smaller portion of the total upgrade cost. This summarises that system-wide electrification is generation driven to be the major factor in the capital cost, whereas the network upgrades are a representation for the grid infrastructure.

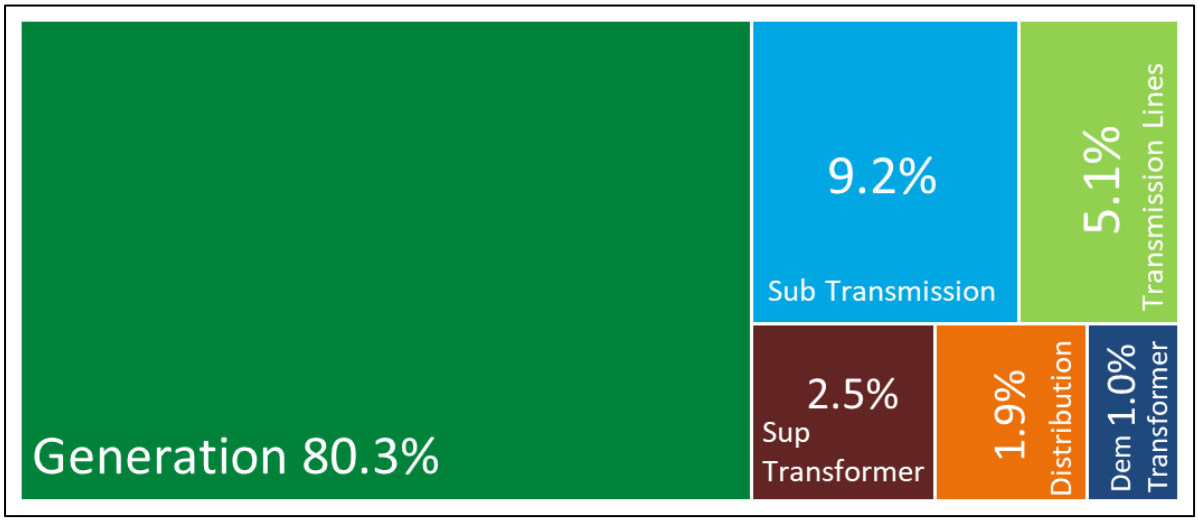


Figure 8.9: Capital cost breakdown of infrastructure upgrades.

The electricity price reduction carries several implications. Firstly, the sequential modelling framework challenges the common assumption that increased electrification would ultimately lead to higher electricity prices. However, with the conditions where renewable generation

technologies are cost-competitive and scalable, electrification can lead to lowering electricity prices compared to the base year for all consumers. In addition, the results highlight the importance of generation technology cost assumptions. The evaluation of generation technologies must consider not only capital investment costs but also variable operating costs, particularly fuel expenses.

Given that the energy modelling framework is versatile, as mentioned in Section A, with cost functions defined as parameters for each technology, capital costs and fuel prices for thermal generation can be adjusted to assess when the balance may shift in favour of thermal generation over renewable alternatives. This flexibility provided by the highly spatial model enables the framework to address broader future research questions through multiple scenario runs. Therefore, the conclusion remains conditional, because the outcome results depend on the interaction of multiple variables and constraints as well as multiple modelling assumptions rather than representing a universal result.

Furthermore, the results from the grid reinforcement modelling demonstrate that a high-level of spatial detail is essential from a strategic planning perspective. The model identifies necessary upgrades in the transmission and POC layers, highlighting critical segments that require attention to relieve bottlenecks and accommodate increased electricity flows. The sequential optimisation approach enables an iterative loop of updating electricity upgrade requirements in response to the quantity of industrial process heat electrification. From this perspective, the framework proves useful for evaluating system structures at a granular level, allowing what if scenarios to be simulated to inform infrastructure planning without exposing stakeholders to real-world risk during project deployment. For example, the model indicates that electrification of industrial process heat, when aligned with renewable generation expansion, does not necessarily compromise electricity affordability or require fundamental grid redesign. Instead, it may accelerate the transition toward a generation mix characterised by lower marginal costs.

8.6.5 Factory Model

Referring to the factory layer, the fuel switching as a result of accounting for the grid upgrade implications remains broadly consistent with the results presented in Chapter 7. However, there is a slight increase in electrification observed in the sequential optimisation framework which is displayed in Table 8.6. The shift is driven by the 10% reduction in the overall electricity price relative to the factory model. The sequential grid optimisation process determines the adjusted electricity price, which reduces the cost of electricity supplied to the industrial sites for electrification. As a result, the electrification technologies such as heat pumps and electrode boilers become slightly more competitive than the previous assumptions where electricity prices in NZD MWh^{-1} would remain the same after new upgrades. Table 8.6 compares the total annual process heat supplied by each technology category under the standalone factory model as well as the integrated factory with a grid model.

The results indicate an increase due to the corresponding decrease in biomass utilisation. The magnitude of change remains small, suggesting that biomass is still cost-economical in most

instances. The outcomes are supported by Figure 8.10 and Figure 8.11, which display fuel-switching results by region and by sector.

Table 8.6: Comparison of fuel-switching results by end-use demand between Chapter 7 and the sequential grid model.

Comparison	**Heat Pumps (TWh yr^{-1})**	**Electric Boilers (TWh yr^{-1})**	**Biomass-based Heating Technology (excluding wood and pulp industries) (TWh yr^{-1})**
Factory Model (TWh yr^{-1})	2.48	1.84	5.48
Factory Model + Grid Model (TWh yr^{-1})	2.61	1.84	5.35
Change (%)	+6%	+6%	-9%

The model also shows that biomass transport distances remain relatively short. Biomass supply is sourced from nearby resource nodes, reflecting the optimisation's preference for minimising transport costs. Since biomass transport cost contributes directly to total system cost, shorter travel distances improve the economic competitiveness of biomass-based heating technologies. This spatial proximity effect reinforces biomass deployment in resource-rich regions while limiting long-distance transport. Figure 8.12 presents the biomass-to-factory connections identified in the optimised solution so as to highlight the spatial structure of biomass allocation strategy. In addition, Figure 8.13 further shows the biomass resource utilisation for industrial process heat by region. It demonstrates that biomass consumption is concentrated in regions where there are strong forestry and agricultural resource presences.

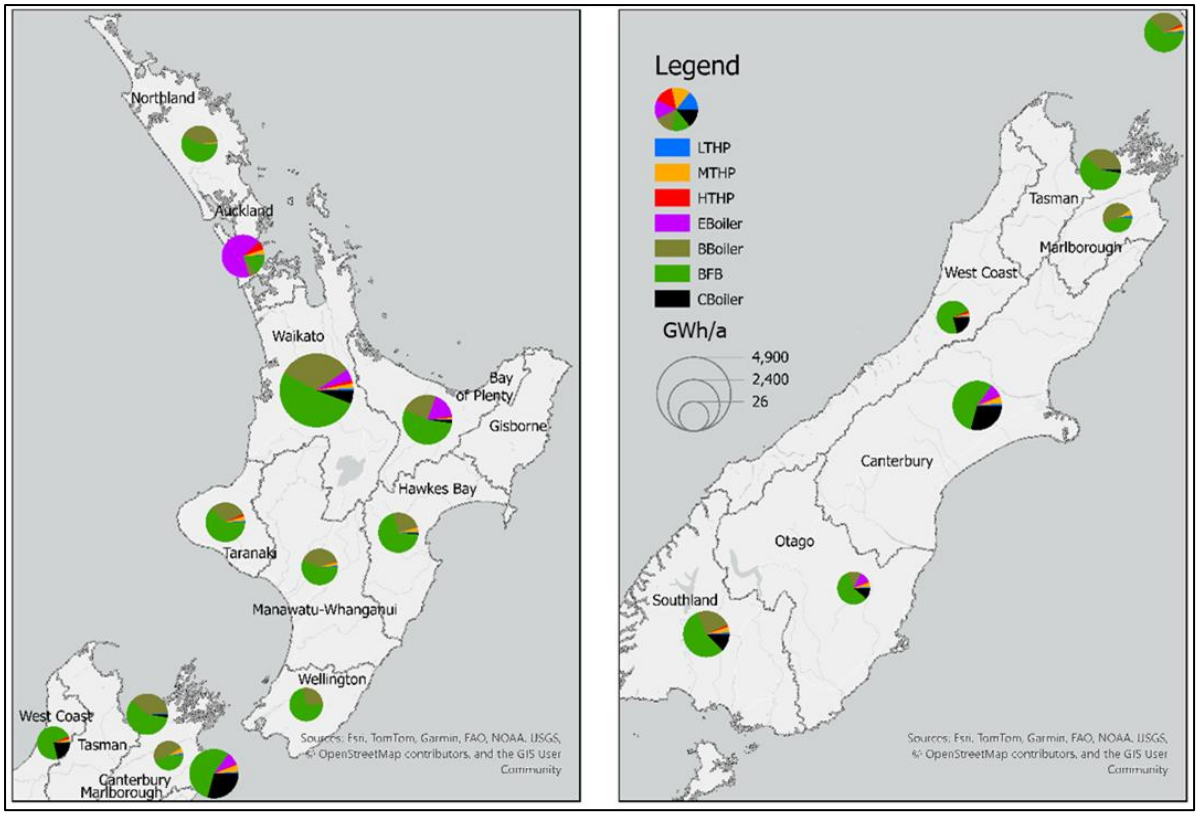


Figure 8.10: Fuel-switching outcomes identified in the optimised solution by region.

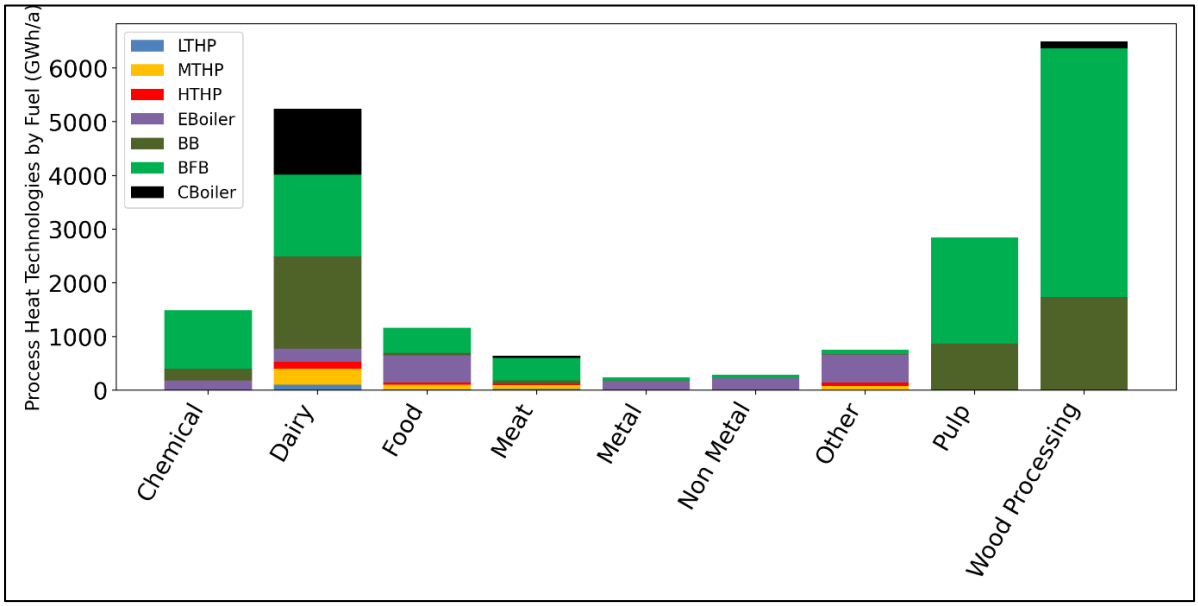


Figure 8.11: Fuel-switching outcomes identified in the optimised solution by sector.

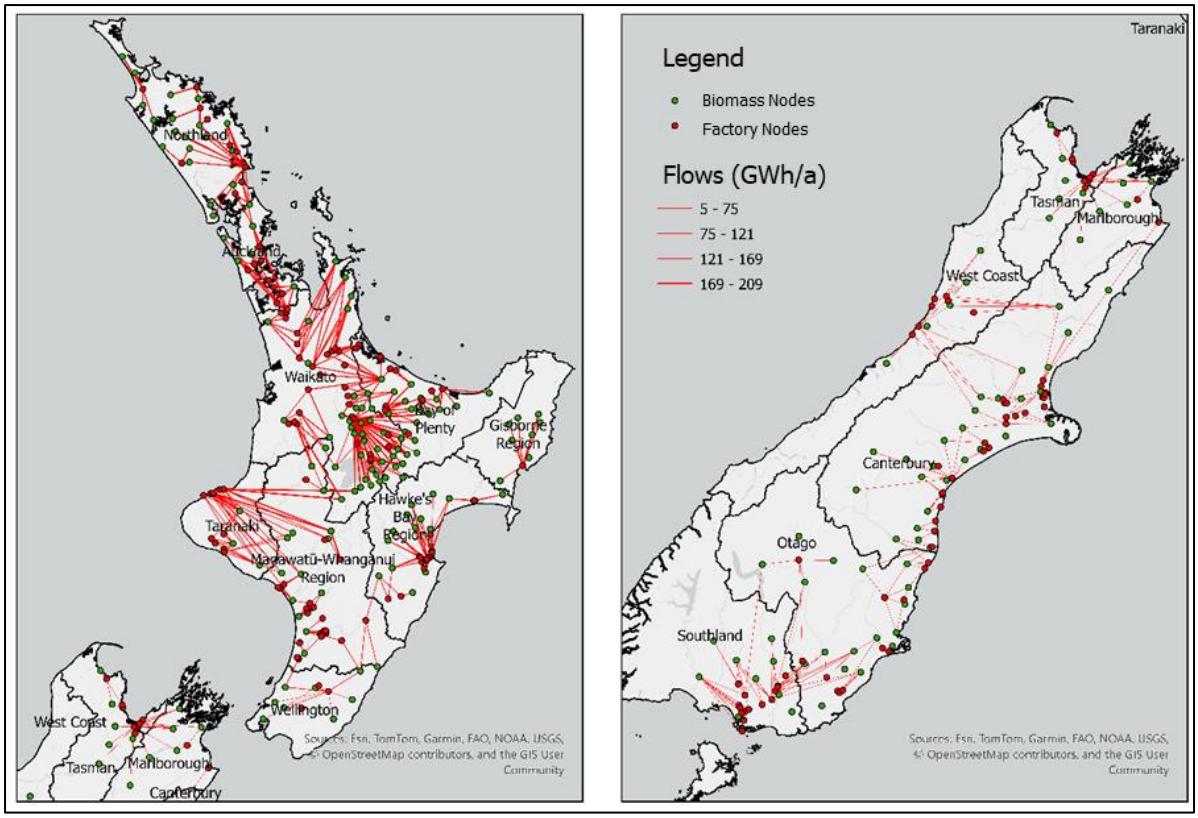


Figure 8.12: Regional spatial distribution of biomass and factory connections established in the optimised solution.

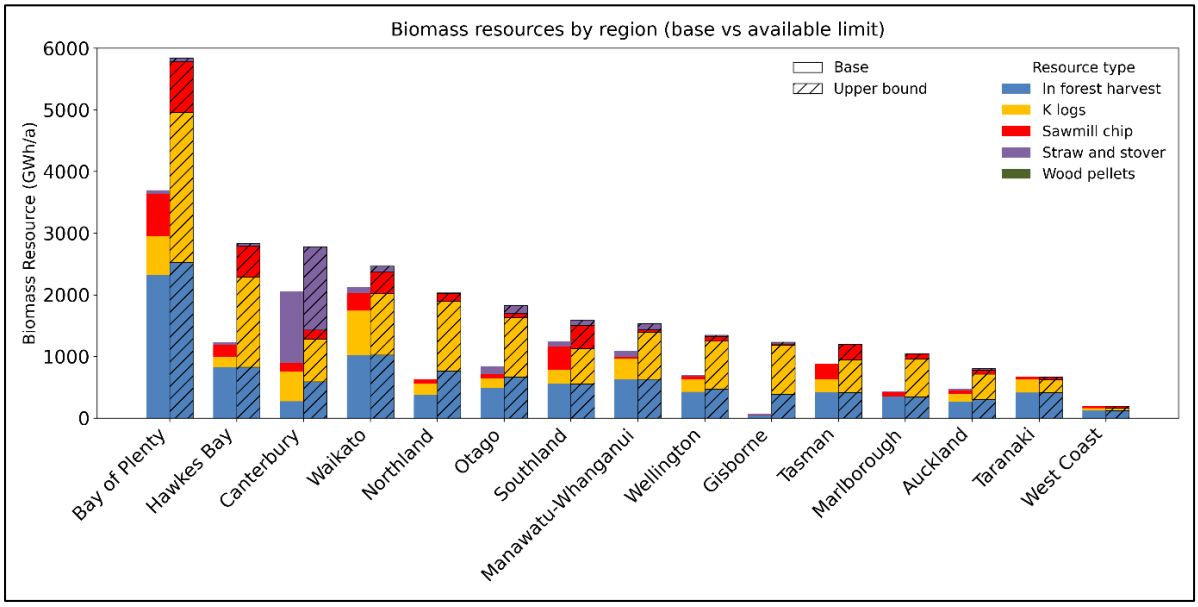


Figure 8.13: Biomass resources used for process heat in industries by region.

8.6.6 Fully Electrified Scenario

An alternative scenario is examined to evaluate the implications of full electrification across all industrial sectors, where all decarbonisable process heat demand is supplied exclusively by heat pumps and electrode boilers. Under this scenario, total electrification increases substantially from 2.81 TWh per annum in the base case to 12.4 TWh per annum.

Figure 8.14(a) and Figure 8.14(b) present the breakdown of energy flows through each electrified heating technology in comparison with the base case. In total, approximately 17.75 TWh of energy is supplied to industrial sites under the fully electrified scenario. The most significant change is the increased deployment of HTHP. This is primarily driven by the high demand requirement to be decarbonised within the 90–140°C temperature range, which accounts for 6.5 TWh of end-use demand, as illustrated in Figure 8.14(c). Given that the total end-use demand is 16.7 TWh, a substantial proportion of process heat demand lies above 90°C, making HTHP a viable and dominant technology in the full electrification scenario.

The increase in electrification leads to an additional 22.9 TWh per annum of electricity demand on the grid. As a result, existing thermal generation units at Huntly (Units 1, 2, 4, and 5) are retained in the model in retrofitted form, while no new thermal generation capacity is commissioned. Renewable technologies dominate capacity expansion. Approximately 5,520 MW of wind capacity is installed, generating 11.31 TWh of electricity annually. The wind capacity is complemented by 423.5 MW of additional hydro capacity and 1,022.4 MW of geothermal capacity, both of which are fully utilised due to resource availability constraints. Together, the electricity generation would increase up to 60 TWh per annum in the full electrification scenario. The electricity generation by technology and by region is presented in Figure 8.15.

Despite the increase in electrification, the model consistently favours renewable electricity generation technologies over biomass-based thermal generation. This trend indicates that biomass is not cost-effective for electricity generation but remains more suitable for direct use in industrial heating applications. One explanation lies in the relative fuel costs: sub-bituminous Indonesian coal is priced at approximately 7.5 NZD GJ^{-1} and natural gas at 9.54 NZD GJ^{-1}, whereas woody biomass ranges from 15.7 NZD GJ^{-1} to 27.7 NZD GJ^{-1} depending on location and resource type.

In summary, the model prioritises lower marginal cost renewable technologies for electricity supply. However, it is also shown that the full electrification scenario has many challenges, particularly to increase electricity generation to 60 TWh per annum in 2037 when the 2022 electricity generation figure is 38.9 TWh per annum.

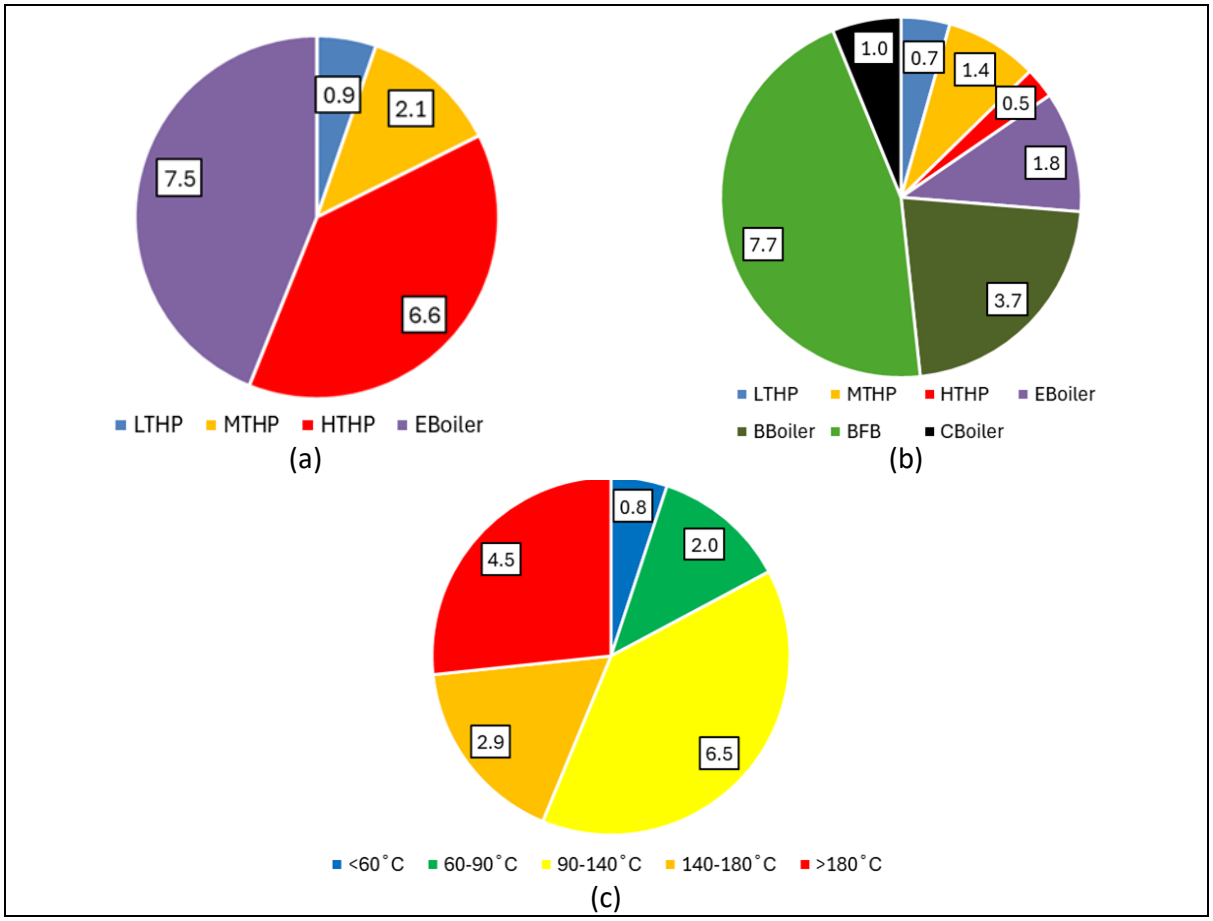


Figure 8.14: (a) Fully electrified scenario-based on industrial end-use demand in TWh yr^{-1}. (b) Base case model in TWh yr^{-1}. (c) Distribution of heating demand across temperature bands TWh yr^{-1}.

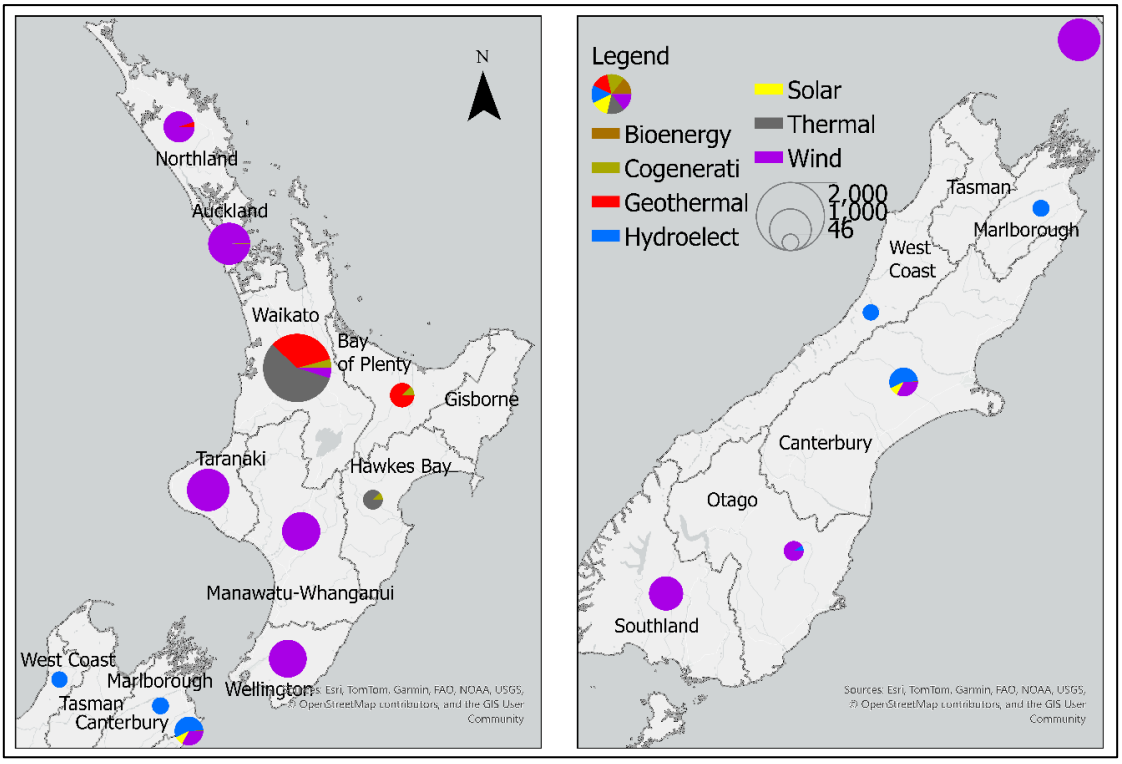


Figure 8.15: Spatial distribution of new proposed generation in MW, including retrofitting thermal and cogeneration generation and commissioning bioenergy capacity for the full electrification scenario.

8.7 Limitations and Conclusions

The sequential optimisation framework which was carried out has lined the factory-level model and the grid reinforcement model. The approach shows how industrial electrification outcomes would translate into electricity system upgrades and how these upgrade costs can be fed back into the electricity sales price through the updated electricity price coefficient $\alpha^{e\ adjusted}$. The optimisation framework is effective in identifying grid reinforcement upgrades, and how are the upgrade costs fed back into the electricity sales price in the factory-level model through the updated electricity price coefficient $\alpha^{e\ adjusted}$. The framework is effective in identifying the cost-optimal configurations for the grid reinforcement needed and also in quantifying the location and sizing required for the associated upgrades as a result of industrial electrification and non-industrial demand increase in the projected year.

One of the further works which can be done is the electricity price adjustment can be done in different levels of spatial detail, corresponding to the line companies within the region. In this chapter, an average $\alpha^{e\ adjusted}$ is applied uniformly across sectors and regions. This can be implemented by varying $\alpha^{e\ adjusted}$ at either the POC level or the territorial authority level. In this way, the electricity price adjustment can reflect the spatial nature of the upgrades undertaken as well as the capability of the modelling framework which works on a site-specific level. This is because by applying several electricity price adjustments in the sales electricity cost calculation, the upgrades in the local distribution level and larger regional or transmission-scale investments can be determined, which would influence the fuel-switching decision for an industrial process site. Such differentiation would allow the model to allocate cost impacts more accurately according to where the infrastructure upgrades occur, rather than applying a single national adjustment.

An important outcome of the sequential modelling framework is the demonstration of a novel implementation approach within the P-graph environment. The decision variable results from the first-stage model are extracted and introduced as fixed inputs in the second-stage formulation. This enables a structured transfer of solution information between optimisation stages, which runs iteratively while preserving superstructure consistency.

Such a pipeline cannot be automated using P-graph Studio alone. Similarly, it is not directly achievable within standard P-graph Python workflows without the development of additional preprocessing and postprocessing tools. In this study, the transition is enabled through a P-graph to Pyomo porting function, combined with a structured Excel output extraction and text file parsing using a regular expression operations approach, as described in Chapter 5. This extended workflow allows systematic extraction of solution structures and their integration into the subsequent optimisation stage, while using a powerful optimisation solver.

Another contribution of the sequential framework is the integration of numerical root finding techniques from the SciPy library to resolve the convergence between two optimisation stages. The sequential structure has an outer loop problem where the electricity price coefficient $\alpha^{e\ adjusted}$ is iteratively updated until both stages converge.

Several key limitations should be noted. The electricity network is modelled under an N security criterion instead of an N-1 contingency standard. This results in grid reinforcement requirements and costs that are likely to represent a lower bound because additional redundancy would be required under stricter reliability planning. On the other hand, as the grid representation captures major reinforcement drivers, not all network components and operational constraints, such as buses, help to manage computational tractability, but it also means that the model focuses on the dominant equipment upgrades rather than a full engineering replication of the system. Therefore, the results should be interpreted as a screening-level indication of upgrade requirements, as the model could show regions with high stress points and exact upgrade points.

The upgrade cost estimates in the chapter are interpreted as a Class 5 level estimate, this is often used for initial project screening for capital expenditure and planning without spending a large amount of time or effort for each project. Class 5 estimates are defined to have an accuracy of -50% on the low bound and +100% on the higher bound of the project costing. Greater accuracy would require more detailed and higher quality input data, including asset condition, local distribution constraints, protection requirements, and site-specific upgrade designs. Nevertheless, the framework demonstrates credible behaviour in the modelling layer, where upgrade decisions can be compared against known corridor constraints and asset characteristics, providing confidence that the model is capturing key structural limits of the network.

Overall, the sequential optimisation framework is effective in reducing the size of the optimisation model while at the same time being proficient in diagnosing grid reinforcement requirements from industrial electrification and translating these requirements into an electricity price adjustment concept. Under the assumptions made, the model indicates that the updated electricity sales price would reduce by 9% across all sectors. This result is conditional based on the cost structure of technologies, reliability criteria, and asset representation. Hence, it should be interpreted as an outcome instead of a prediction. Future works include improving data quality, such as reliability criteria, expanding spatial reporting of $\alpha^{e\ adjusted}$ so that the electricity price can be more attributed to local upgrades.

While the sequential optimisation approach in this chapter provides insights into system configurations under deterministic assumptions, it does not account for uncertainty in key modelling parameters. In reality, modelling parameters such as biomass availability, electricity price changes, and biomass resource costs are subject to large variability, leading to changes in system configuration and investment decisions. Therefore, Chapter 9 extends the factory model analysis by incorporating a Monte Carlo-based uncertainty framework, which provides a probabilistic assessment of system outcomes and robust testing of decarbonisation pathways.

Chapter 9: Monte Carlo Uncertainty Analysis

9.1 Introduction

Chapter 9 extends the deterministic multiperiod model by evaluating how uncertainty in several key input parameters affects the robustness of the least-cost solution. Energy modelling systems are inherently complex and subject to unpredictability. This makes long-term projections difficult to represent using fixed assumptions or a single baseline set of parameter values. While the deterministic optimisation approaches presented in Chapters 7 and 8 identify least-cost solutions under deterministic input conditions, real-world energy planning must account for variability and unpredictability across both supply and demand. The Monte Carlo framework complements the multiperiod and sequential optimisation formulations by introducing stochastic variation in parameters that cannot reasonably be treated as fixed. This provides the basis for an uncertainty analysis in which Monte Carlo techniques are used to assess fuel-switching pathways across a wide range of plausible scenarios. Hence, the chapter establishes the robustness of Chapter 7 and 8 findings to uncertainty in key parameters, which completes the triangulated evidence referenced in the conclusions.

This chapter presents the Monte Carlo method within the energy modelling framework to evaluate the uncertainty and variability of the input parameters. Instead of relying on a single baseline dataset, the model is solved repeatedly under randomly sampled inputs. Each simulation run presents one possible scenario within the uncertainty range, it allows the feasible pathways and cost outcomes to be assessed statistically. This approach highlights that the most relevant configuration is not always the solution obtained under a single deterministic case but instead the most robust configuration that remains feasible and cost-effective across many operating conditions. By incorporating repeated sampling, the modelling framework provides additional flexibility for evaluating the sensitivity of optimal solutions.

A key motivation for applying Monte-Carlo method is contributed by the computational limitations of the P-graph framework. Although P-graph algorithms are capable of identifying multiple near-optimal (n-best) solutions, this feature becomes increasingly difficult to apply for larger sized problems, as each branch and candidate structure must be evaluated explicitly in the search for the optimal branch. For high-dimensional complex energy system models, the size of the solution space would be impractical to create combinatorial decisions. Therefore, Monte-Carlo method is an alternative to that for exploring variability in outcomes without requiring full enumeration of all structural alternatives.

In this chapter, the same case study is applied. Random variation is introduced through uncertain inputs which are changes in electricity demand, electricity prices and biomass resource availability. In the case study, results are interpreted from the risk-informed interpretations for the process heat transition in industries. A 4-hour temporal resolution is used to demonstrate the applicability of Monte Carlo methods in capturing uncertainty at higher operational detail within the energy modelling framework.

9.2 Monte Carlo Concept

The Monte Carlo method is applied in this study as a reliability-based uncertainty evaluation tool that can be integrated directly into the Python modelling workflow. Instead of treating all parameters as fixed deterministic inputs, the Monte Carlo method enables variables to be sampled from a set of predefined probability distributions. The optimisation model is repeatedly solved under the same superstructure under varied randomly generated inputs. The workflow integrates directly with the Python modelling environment because all model inputs are treated as exogenous parameters obtained from the energy system model, leading to no implications on the model structure as the distributions are conducted beforehand. Stochastic parameters are generated externally before passing into the optimisation formulation. This construction of the module provides a degree of flexibility for model execution by allowing users to specify:

(i) the maximum allowable number of simulation runs
(ii) a tolerance value which enforces an automated stopping rule on the repetitive simulation runs until key outputs converge.

These features allow the uncertainty analyses which are carried out to be scaled depending on the computational limits and modelling requirements.

The first approach of stopping the Monte Carlo is the fixed sample-size procedure. The Monte Carlo analysis is conducted using a maximum of $j_{max} = 1000$ simulation runs. Each run represents five varying sampling parameters obtained from a set indexed by $\omega \in \Omega$. Therefore, in every run, the optimisation model is solved independently between runs with the varying inputs. In order to ensure reproducibility and traceability of the thesis's results, random seeds are deliberately defined.

A key question in the Monte Carlo uncertainty analysis is to determine when the simulation run stops and how many samples are needed. Apart from confidence interval-based stopping rule applications where the cumulative running mean is close to the expected sample value, the acceptable shifting convergence band rule (ASCBR) can also be used to determine the Monte Carlo sample size. While both methods can be used, the ASCBR method is applied in this model. This is because the confidence interval-based stopping rule has a higher computational complexity, as the sample variance has to be repeatedly calculated during each simulation run. The computational efficiency is a particularly important aspect to consider in this energy modelling framework due to the size and complexity of the problem.

On the other hand, ASCBR only compares means with band limits, where the running mean must stay within a band of width 2ε with a number of iterations for convergence f^*. A band of width 2ε is to account for the upper and lower limits of the band in search for the convergence run. A strong heuristic of determining the value of f^* is to be greater than 50 (Ata, 2007). Figure 9.1 illustrates the idea behind the stopping rule being enforced in the model.

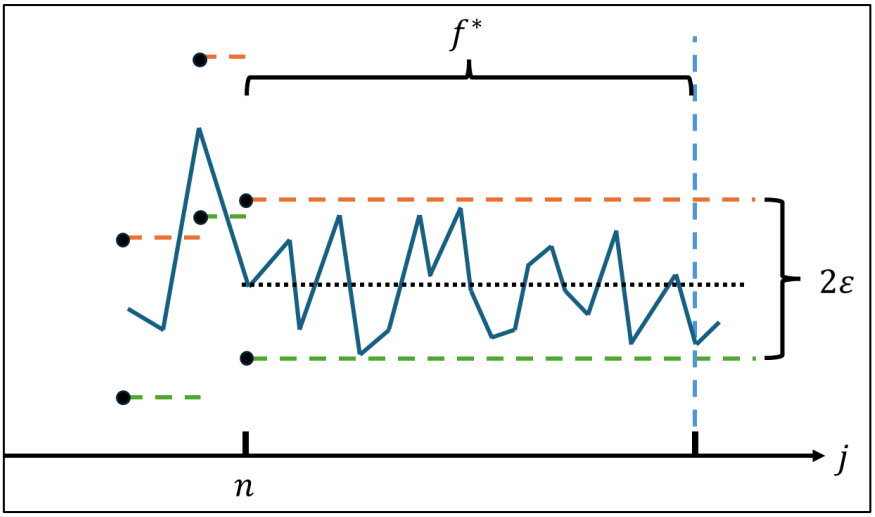


Figure 9.1: Illustrative example of convergence band for each Monte Carlo run.

The Monte Carlo analysis is equipped with a function where it would terminate its operation when the run reaches $j = 1000$ or a convergence criterion ε of 1% is met for the normalised cumulative mean value $\bar{x}_j$. This means that the $\bar{x}_j$ would achieve a full bandwidth of 2% in reaching the expected value $\mu = \langle X \rangle$ A running cumulative normalised mean is expressed as:

$$\bar{x}_j = \frac{1}{j}\sum_{i=1}^{j} x_i \qquad (9.1)$$

where n is the number of simulation runs which are completed.

The sample variance or mean squares of the distances s_j^2 from the sample mean, is denoted as:

$$s_j^2 = \frac{1}{j-1}\sum_{i=1}^{j}\left(x_i - \bar{x}_j\right)^2 \qquad (9.2)$$

A Monte Carlo process is regarded as strictly stationary because the sample mean obtained from multiple MC simulations converges in probability to the expected value $\mu = \langle X \rangle$. Hence, the expression can be described as:

$$\lim_{j\to\infty} Pr(|\bar{x}_j - \mu| < \varepsilon) \to 1 \qquad (9.3)$$

As a summary, the protocol and setup for the Monte Carlo analysis framework is reproducible and transparent so that the number of simulation runs is sufficient to obtain key indicators without excess computational burden.

9.3 Temporal Structure and Stochastic Parameters

In this section, the stochastic input generation methods are discussed in the Monte Carlo uncertainty analysis. The sampling framework is defined to ensure the transparency and reproducibility of the thesis. The Monte Carlo uncertainty analysis in this chapter is implemented using a different temporal structure from the deterministic optimisation framework. To maintain

computational tractability across repeated runs, the model is formulated using a 4-hour temporal resolution over a representative daily horizon. This higher operational resolution enables the assessment of short-term variability while maintaining computational tractability across repeated simulation runs.

Selected parameters are treated as stochastic inputs within this temporal resolution. Electricity demand and wholesale electricity price data are derived from the Electricity Authority database. On the fuel side, biomass resource limits and biomass prices which are regionally dependent are sourced from Hall (2022). A uniform random sampling method is applied to assess the uncertainty of biomass for fuel-switching in heating technologies in industries. On the other hand, capacity factor for intermittent solar and wind generation at an hourly resolution is obtained from the Renewables.ninja reanalysis weather data.

Although five uncertain inputs were defined in the Monte Carlo structure, only three were active in the factory model optimisation stage, which are examined here. Two additional inputs, which are the electricity demand and capacity factor, were included to preserve compatibility with a later model extension. Both parameters were not operationalised in the present analysis and therefore had no effect on the optimisation results reported in this thesis. These two are only needed when a grid reinforcement model is implemented within the multistage framework. Their inclusion here serves only to preserve the structural consistency of the modelling framework. As the sequential optimisation was already performed in the previous chapter, the focus of this chapter is limited to demonstrating the Monte Carlo methodology.

9.3.1 Electricity Prices

Electricity price uncertainty is modelled by the average wholesale electricity price to reflect the long-term pricing strategy of FPVV contract pricings commonly applied to large-scale industrial users. A historical window of eleven years of annual average whole electricity prices from 2013 to 2023 are used as the sampling pool. The historical price distribution has a mean μ_P of 178.71 NZD MWh^{-1} and a standard deviation σ_P of 13.56 NZD MWh^{-1}.

The Monte Carlo run ω, the wholesale price is sample with a normal distribution formula:

$$P^{\omega}_{wholesale} \sim \mathrm{N}(\mu_P, \sigma_P^2) \qquad (9.4)$$

The sampled wholesale electricity price is converted into an electricity sales tariff using the FPVV formulation developed by the Ahuora Energy Centre. The result of this sampling strategy produces samples for all 4-hour blocks within each Monte Carlo run. The approach provides better consistency in results, as the FPVV behaviour is more realistic than short-term spot price volatility.

9.3.2 Electricity Demand

On the other hand, electricity demand uncertainty is modelled at 4-hour resolution. The historical demand data from 2022 to 2024 are used to construct baseline demand profiles. Aggregating three years of hourly data provides approximately 26,280 number of datasets per time block for each POC. This provides a robust sample size for all 166 recorded electricity demands in the

Electricity Authority Database. Instead of applying independent random sampling at each time step, the change in demand or the demand trajectories are generated sequentially to preserve the coherence between time blocks.

The midnight demand level (Hour 0) is treated as an anchor point for each simulated day. Each Monte Carlo run ω, the formulation is expressed as:

$$D_0^{\omega} \sim \mathrm{N}\left(\mu_Q, \sigma_Q^2\right) \quad (9.5)$$

where μ_Q denotes the historic mean and σ_Q^2 represents the standard deviation of the demand at midnight. In this study, negative sampled values are set at zero to ensure physical feasibility of the system.

$$D_0^{\omega} = \max(0, D_0^{\omega}) \quad (9.6)$$

A single random draw is applied across all POCs for each simulation run. This is because the electricity demands are correlated with one another at the same temporal time, preventing unrealistic divergence between nodes and preserving the structural coherence of the national load profile.

This is followed by sampling of the subsequent 4-hour demand values, which are generated by sampling demand rather than absolute demand levels. This approach ensures that there is temporal continuity across time steps, reducing unnecessary volatility and also preserving realistic intraday load patterns across all POCs. The 4-hourly demand level is then updated sequentially with the following expression:

$$D_h^{\omega} = \max(0, D_{h-1}^{\omega} + \Delta D_h^{\omega}) \quad (9.7)$$

9.3.3 Biomass Resource Cost

Biomass resource costs are treated as uncertain parameters within the Monte Carlo framework. Unlike electricity prices and demands where large historical sample data are available and recorded, biomass material costs in NZD t^{-1} are given a range by Hall (2022) and EECA (2025c). This is attributed particularly to the moisture content of the resource or market demand, leading to the change in net calorific value of the given fuel type when used as process heat. Biomass material cost has distinct biomass resource costs depending on geographical constraints for the North Island and South Island. For each biomass type, a symmetric uncertainty band of 20% is applied across the biomass cost to represent variability in market prices.

Although cost data for some biomass resources are available, particularly for established wood products traded in domestic and export markets, these prices are typically updated quarterly because such products are treated as commodities. However, comparable price data are not available for the wood resources considered in this model, as they generally consist of low-quality residues that are not suitable for further processing or commercial markets. Therefore, a uniform distribution with defined upper and lower bounds is applied to represent the range of biomass prices in the model. This approach is a better strategy when dealing with uncertainty in biomass

resource costs. The upper and lower bounds for the biomass raw material price are constructed as follows:

$$C^{low} = C^{base}(1-\delta), \qquad C^{high} = C^{base}(1+\delta) \tag{9.8}$$

where $\delta = 0.20$ which is the uncertainty level.

For every Monte Carlo run, the costs of each biomass type are sampled independently from a uniform distribution bounded by these lower and upper limits:

$$C^{\omega} \sim U(C^{low}, C^{high}) \tag{9.9}$$

The sampling is performed separately for the North Island and South Island as well as the biomass types, preserving regional differentiation and fuel type affecting input cost.

9.3.4 Biomass Recoverability

Uncertainty in biomass availability is represented through recoverability factors derived from two reference levels, denoted by L1 and L2. These two represent the recoverability factors for each biomass type based on the gross volume of the resource, where L1 is the higher recoverability factor scenario and L2 is the lower recoverability factor scenario introduced by the Hall (2022). Rather than fixing the recoverability of biomass availability for each biomass node to be a discrete scenario, recoverability for each biomass category is sampled from a uniform distribution bounded by the two levels, which can be represented as the following:

$$\lambda^{\omega} \sim U(L1, L2) \tag{9.10}$$

This approach allows each simulation run to reflect a different feasible biomass supply level within a set range. Therefore, by applying a uniform distribution for biomass resource availability and price, the Monte Carlo run would provide a better understanding of returning solution structures which are robust and less susceptible to fluctuations. Due to the lack of data quantifying biomass resource variability across monthly, weekly, or daily timescales, it is assumed that the biomass available within a daily horizon is uniformly distributed based on the annual recoverability values generated from the random sampling for each biomass node.

9.3.5 Capacity Factor for Wind and Solar

Capacity factors for intermittent energy technologies such as wind and solar, which are derived from the Renewables.ninja database using reanalysis weather data. A representative historical year is selected as the reference dataset, and hourly capacity factor time series are obtained for wind and solar generation sites. As the data is simulated based on weather data, it is bias-corrected and treated with noise. Hence, the absolute data does not require the study of the change, and the absolute value is good enough.

To incorporate the capacity factor for the intermittent technologies within the Monte Carlo framework, capacity factors are treated as stochastic variables. For each technology type and

location, the historical hourly time series is recorded and aggregated to 4-hourly time blocks to be used to calculate the mean and standard deviation of the capacity factors.

The Monte Carlo run is simulated; 4-hourly capacity factors are generated using a normal distribution:

$$cf^{\omega} \sim N(\mu^{cf}, \sigma^{cf}) \quad (9.11)$$

To ensure the physical feasibility, the capacity factor is bounded within:

$$0 \leq cf^{\omega} \leq 1 \quad (9.12)$$

This approach preserves the variability in historical weather generation, allowing stochastic variation across simulation runs. By sampling renewable capacity factors in this method, the Monte Carlo framework reflects interannual weather variability and the impact on electricity generation volume, thereby improving the robustness of electrification pathway assessment.

As a summary for the random sampling of the model parameters, Table 9.1 shows the parameters implemented, and the following independent assumptions made are as follows:

- Electricity price and demand are sampled independently.
- Biomass availability and biomass resource price are sampled independently from electricity inputs.
- Capacity factors are sampled independently across simulation runs but preserve the intra-day structure derived from the historical dataset.
- As operational data for industrial heat demand are not publicly available within the model dataset, the demand profiles are derived from generic sectoral load profiles. These profiles have already been noise-adjusted, which means that there is no reliable standard deviation available for random sampling. Therefore, this parameter is not treated as a stochastic variable in the Monte Carlo uncertainty analysis.

Table 9.1: Stochastic parameters implemented in the Monte Carlo uncertainty analysis.

Parameter category	Variable sampled	Sampling method	Notes
Electricity price	Annual wholesale price	Random draw from historical distribution from 2013 to 2023	Converted into FPVV tariff
Biomass availability	Recoverability factor	Uniform distribution between L1 and L2 recoverability factor scenarios	Controls regional biomass limits
Biomass price	Regional biomass price	Uniform distribution on dollars per energy range which differs due to carbon content and moisture content	Geographical split still applies
Electricity demand	POC demand scaling	Random baseline trajectory with proportional variation	Applied across 4-hour blocks
Capacity Factor	Hourly capacity factor	Normal distribution	Applied across 4-hour blocks

9.4 Model Formulation

9.4.1 Objective Function

In the Monte Carlo method, the objective function inside each optimisation run stays the same as the single deterministic formulation. The only variation is the electricity price, biomass price and biomass availability that are sampled. Therefore, all formulations remain the same, as the parameters are different for each simulation run. Hence, the objective function is shown as:

$$\min CC^{\omega}_{Resource} + CC^{\omega}_{Op} - CC^{\omega}_{Product} \quad (9.13)$$

where $\omega \in \Omega$ denotes a Monte Carlo run.

9.4.2 Constraints

This section presents the constraint formulation for industrial process heat supply, with particular emphasis on the inclusion of short-term thermal storage. The equations introduced here follow the same deterministic process heat framework developed in Chapter 7, since the Monte Carlo analysis only introduces uncertainty through variations in the input parameters. While these uncertain parameters are sampled across Monte Carlo runs, the underlying structure of the process heat balance, technology supply, and feasibility constraints remains unchanged.

Process heat

Thermal storage is not considered suitable for long-duration applications when the model is solved at a monthly resolution. Therefore, it is excluded from the system-wide optimisation presented in Chapter 7 and Chapter 8. However, when the temporal resolution is changed to a 4-hour interval, industrial facilities may realistically deploy short-term thermal storage at the plant level to buffer intraday fluctuations in heat demand and supply. Thermal storage is introduced as an additional flexibility option within the process heat formulation. Demand is supplied directly by heat technologies or discharge from thermal storage. The expression can be denoted as:

$$D_{l,t,q'} = \sum_{q \in Q} \sum_{p \in P} H^{2}_{l,t,p,q,q'} + Dis_{l,t,q'} \quad (9.14)$$

Thermal Storage

Thermal storage is represented across discrete temperature storage bands q''. In this thesis, only low-temperature hot water storage is considered. Therefore, the storage temperature set is restricted to $Q'' = \{q''_1, q''_2\}$ corresponding to the temperature level. The concept of hot water storage with associated variables is shown in Figure 9.2. The state of charge evolves over time as charging increases stored heat, while discharging reduces it:

$$St_{l,t,q''} = St_{l,t-1,q''} - \sum_{q' \in Q} Dis_{l,t,q'',q'} + \sum_{q' \in Q} Ch_{l,t,q',q''} \quad (9.15)$$

The storage capacity limits are imposed as:

$$St_{l,t,q''} \leq Cap^{st}_{l,q''} \Delta t \quad (9.16)$$

The non-negativity constraints ensure the operation is physically feasible.

$$Ch_{l,t,q',q''} \geq 0, \quad Dis_{l,t,q'',q'} \geq 0, \quad St_{l,t,q''} \geq 0 \tag{9.17}$$

Storage Discharge Aggregation

Discharge delivered to a particular demand temperature band is obtained by summing discharges originating from all storage temperature bands:

$$Dis_{l,t,q'} = \sum_{q'' \in Q} Dis_{l,t,q'',q'} \tag{9.18}$$

9.4.3 Temperature Feasibility

A binary feasibility matrix $M^{ht,st}_{q'q''} \in \{0,1\}$ is added to ensure that heat can be stored into the storage temperature levels to hot water temperature storage levels, $q'' = 1, 2$. The total charging flow is set as a non-zero continuous variable.

$$H^3_{l,t,p,q,q'q''} \leq M^{ht,st}_{q'q''} H^1_{l,t,p,q,q'} \tag{9.19}$$

$$H^3_{l,t,p,q,q',q''} \geq 0 \tag{9.20}$$

The total charging flow into storage can be represented as:

$$\sum_{p \in P} \sum_{q \in Q} H^3_{l,t,p,q,q',q''} = Ch_{l,t,q',q''} \tag{9.21}$$

The routed charging flows are constrained to sum exactly to the heat stream allocated for storage charging to prevent double counting of charged energy.

$$\sum_{q''} H^3_{l,t,p,q,q'q''} = H^1_{l,t,p,q,q'} \tag{9.22}$$

The heat generated by each technology is split between direct demand supply and charging into storage. Split stream because $H^{sup}_{l,t,p,q,q'}$ is generated from heat technology.

$$H^2_{l,t,p,q,q'} + H^1_{l,t,p,q,q'} = H^{sup}_{l,t,p,q,q'} \tag{9.23}$$

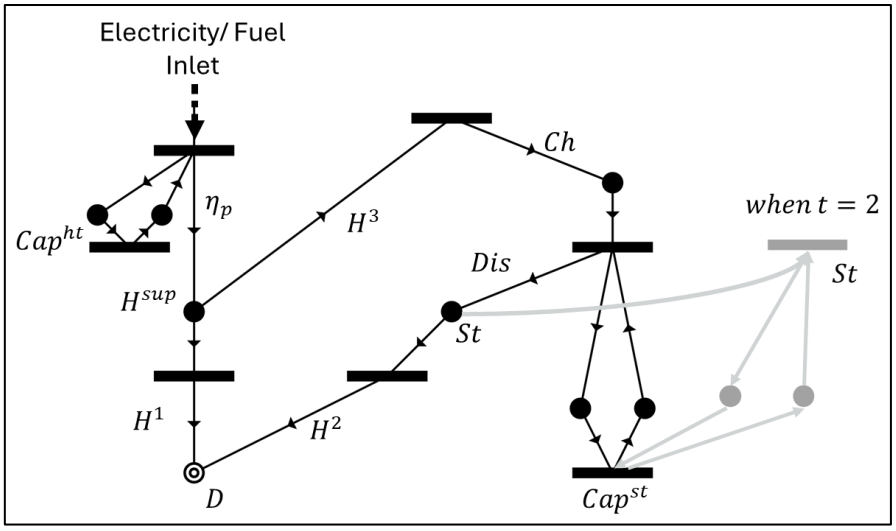


Figure 9.2: Hot water storage design within the P-graph framework, $|T| = 1$.

9.5 Output and Reporting

As each Monte Carlo iteration is solved individually, the computational burden is considerable. Each simulation run requires approximately 800 seconds to reach an optimal solution while maintaining the same solver settings adopted in the previous chapter. This further highlights the importance of applying a stopping rule, because excessive simulation runs would contribute a higher computational cost.

Based on the resulting distribution, the mean total system cost is NZD 10,562,259, whereas the median value is NZD 10,493,182. Referring to Figure 9.3, the distribution of total system cost is bimodal, indicating the presence of two distinct clusters of solutions. This bimodality is caused by two structurally different process heat configurations identified by the model: one that is more biomass-oriented and another that is more electricity-oriented. The existence of these two configurations suggests that the model does not converge towards a single dominant pathway under uncertainty. Instead, it reveals two competing classes of feasible system configurations.

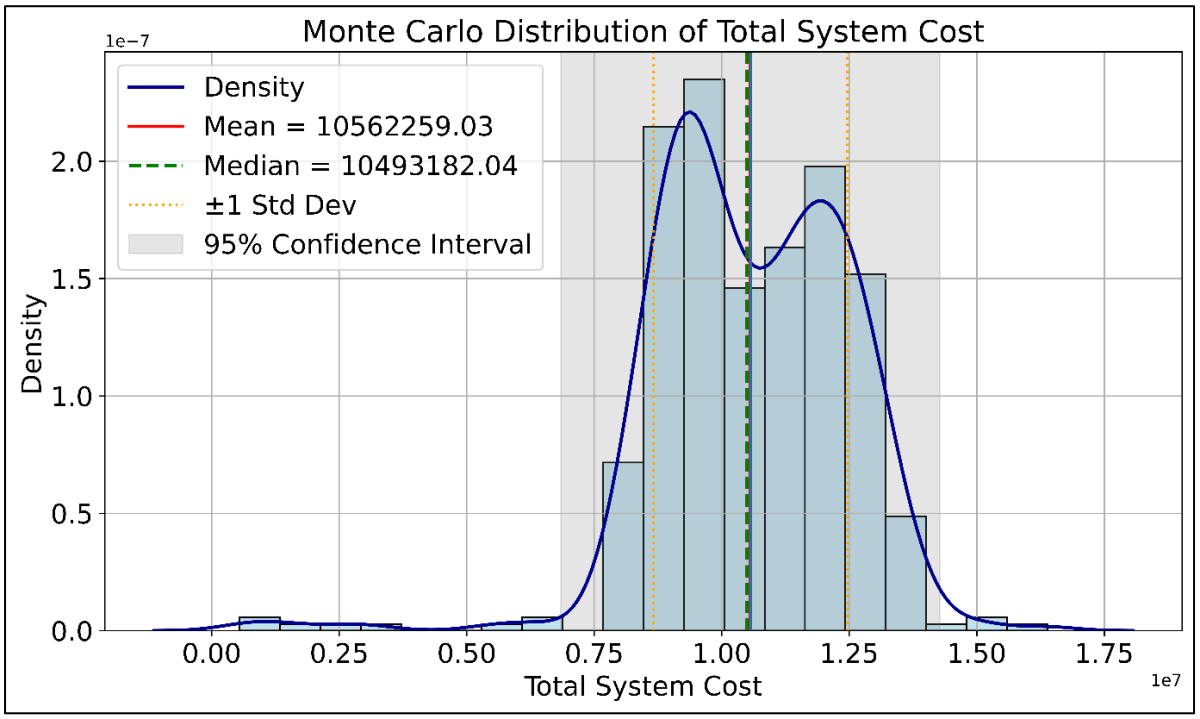


Figure 9.3: Monte Carlo probability density function plot for total system cost.

Based on the running mean and convergence band in Figure 9.4, they indicate that convergence is achieved after 134 simulations under a tolerance of 1%. The lower convergence band is NZD 10,457,205 and the upper convergence band is NZD 10,668,450. The behaviour of the running mean is consistent with the bimodal nature of the Monte Carlo distribution. The cumulative mean fluctuates upward and downward while the convergence bands are updated before stabilising. Once f^* is reached, additional iterations are no longer needed for satisfying the stopping criterion, although it is still useful for observing the subsequent behaviour of the running mean.

Figure 9.5 presents the consecutive count under the ASCBR criterion, further validates the convergence of the Monte Carlo runs. The value of z_j is being reset to 0 when the running mean

falls outside the convergence band. As shown in the figure, the count eventually reaches 50, which surpasses the required f^* value at the 134th simulation run. Before reaching the convergence, there is an earlier peak with $z_j = 31$ before being reset at $z_j = 84$. That indicates that the running mean was close to approaching the expected value before the final convergence was achieved if the tolerance is increased.

The supporting Figure 9.6 displays the running standard deviation after each iteration run is made. The figure shows that the running standard deviation plateaus after approximately 150 iterations. This suggests that the spread of the simulated results becomes stable well before the stopping rule is triggered. Upon reaching the convergence point, the running standard deviation is NZD 1,841,405. This shows that although the running mean continues to be updated after every subsequent iteration, the overall variability of the Monte Carlo outcomes becomes fairly consistent at a much earlier stage.

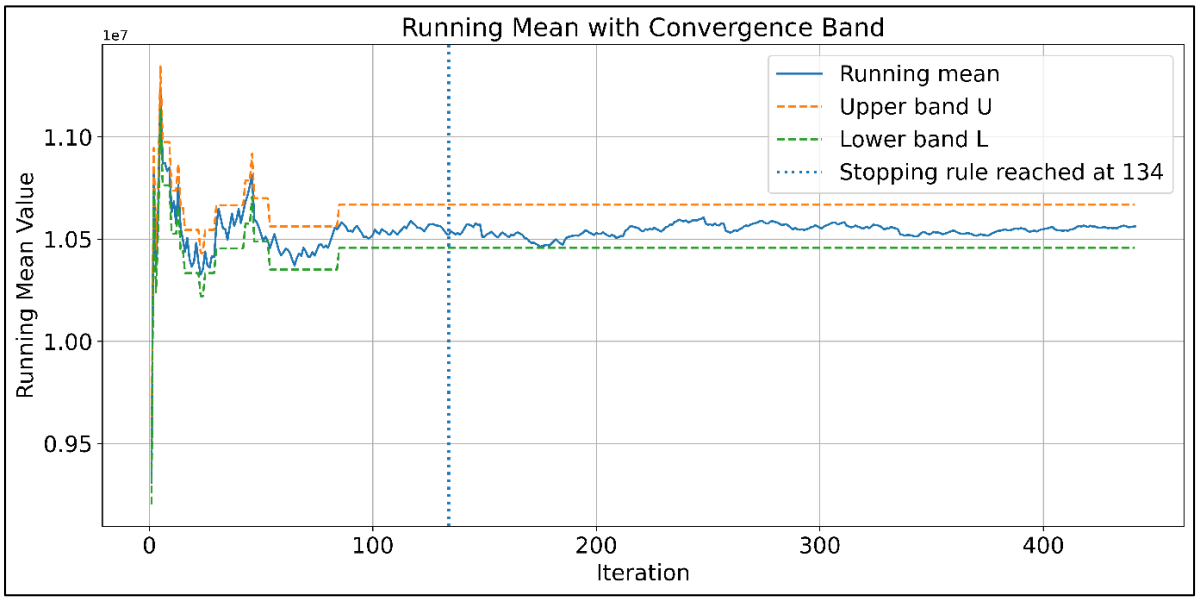


Figure 9.4: Cumulative running mean in a convergence plot.

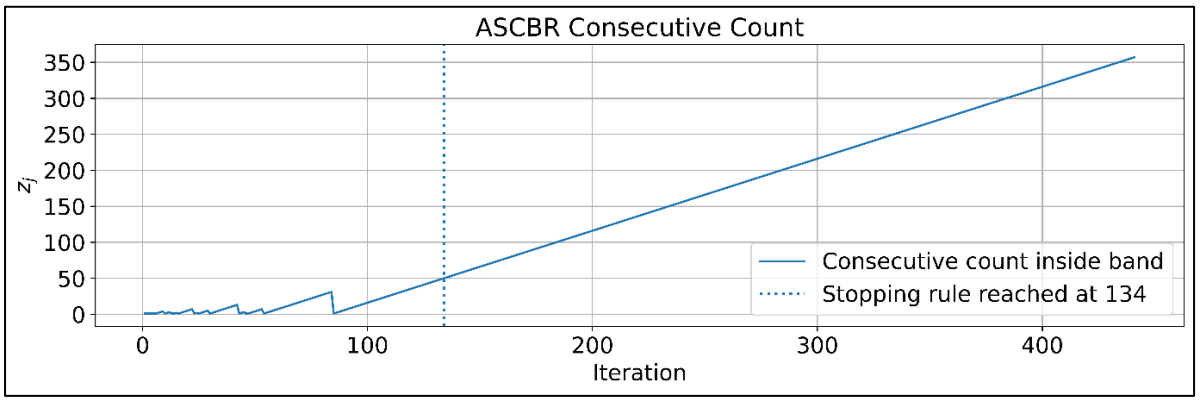


Figure 9.5: Plot of ASCBR consecutive count of random variable z_j.

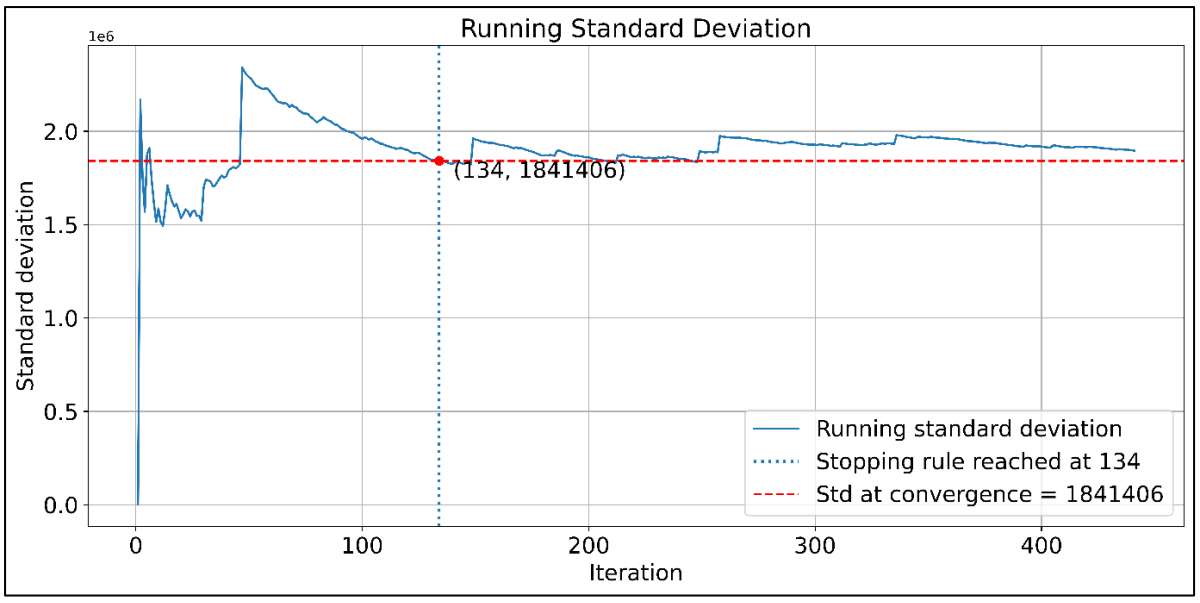


Figure 9.6: Plot of running standard deviation across iteration runs.

Results presentation

The energy modelling framework allows simulation results from all Monte Carlo runs to be extracted and analysed from several perspectives. This is because the purpose of having Monte Carlo runs is to gain insight on what the system behaviours are on a 4-hourly interval. The model is capable of presenting both system-level and individual node-level results. The outcomes are primarily presented using box plots so as to summarise the statistical distribution of different technologies, fuel use and biomass utilisation across all Monte Carlo runs.

By studying from the system level, the aggregated configuration of industrial process heat technologies which includes the distributed installed capacity for heating, as well as the application of biomass resources across the country. This technique provides an overall outlook in terms of fuel-switching in the factory model as well as the associated uncertainties originated by the random sampling approach.

Moreover, the modelling framework allows the results to be critically analysed at an individual factory and biomass supply node level. Similarly, box plots will be used to represent the interactions and behaviour of the fuel-switching technologies and biomass usability under stochastic parameter conditions.

System Level View

From the system-level results, electrified heating technologies play a significant role in the deployment of process heat supply capacity as shown in Figure 9.7. Among all three heat pump technologies, low-temperature heat pumps (LTHP) show a median capacity of 376.6 MW and a mean capacity of 36.6 MW respectively. Medium-temperature heat pumps (MTHP) contribute the largest share among the heat pump technologies, with a median of 507.1 MW and a mean of 345.1 MW, partly due to the high use of MTHP to store hot water as well as to be cascaded down

to low-temperature requirements, making the use of LTHP in the model scenario relatively redundant. The HTHP has an interquartile range of approximately 196.0 MW to 228.8 MW.

A wider uncertainty in fuel-switching uptake can be observed in the box plots for electric boilers, biomass boilers, and bubbling fluidised bed heating technologies. This result indicates a high degree of variability in technology deployment at higher temperature levels, leading to competition between the boiler technologies. This variability reflects the sensitivity of these technologies to fluctuations in electricity prices and biomass availability across Monte Carlo runs. Under certain conditions, lower electricity prices favour electrified technologies, while in other cases greater biomass availability encourages biomass-based heating technologies.

Thermal storage technologies show very limited deployment in the results. Low-temperature and medium-temperature thermal storage capacities remain close to 0 across most simulation runs, with total installed capacity across the system being approximately 1 MW or less. Hence, the model has concluded that thermal storage at a factory level is not economically attractive under the current model assumptions. This is primarily due to the relative cost competitiveness of electrification or biomass-based heat supply. However, this result should be interpreted carefully, as the current factory-level model determines the most economical scenario for the system level of the factory and does not explicitly represent energy security considerations, which may increase the attractiveness of thermal storage in real-world applications.

In terms of the system view of biomass resources, each resource type is utilised within the model. Figure 9.8 indicates that biomass plays a relevant role across the energy flows associated with the biomass node, with most resources being consumed by nearby factories with biomass-based heating technology. The results suggest that biomass utilisation is strongly influenced by location and transport distance, as biomass resources are generally consumed at the closest feasible industrial sites.

The model also shows that some biomass resource nodes remain unused despite having available supply. This typically occurs when the biomass resource nodes are located too far from industrial demand centres, or when electrification prices are sufficiently high enough to reduce the share for biomass-based heating technologies. Spatial distribution also plays a pivotal role in process heat decarbonisation. For example, in forest harvest residues and k logs are more uniformly distributed across the country where commercial forests are present, as their resource availability is independent of the location of industrial sites. On the other hand, sawmill chips are concentrated in regions with significant wood processing activity. In contrast, the Canterbury region, which houses a significant portion of major South Island industrial sites, boasts a higher abundance of straw and stover resources.

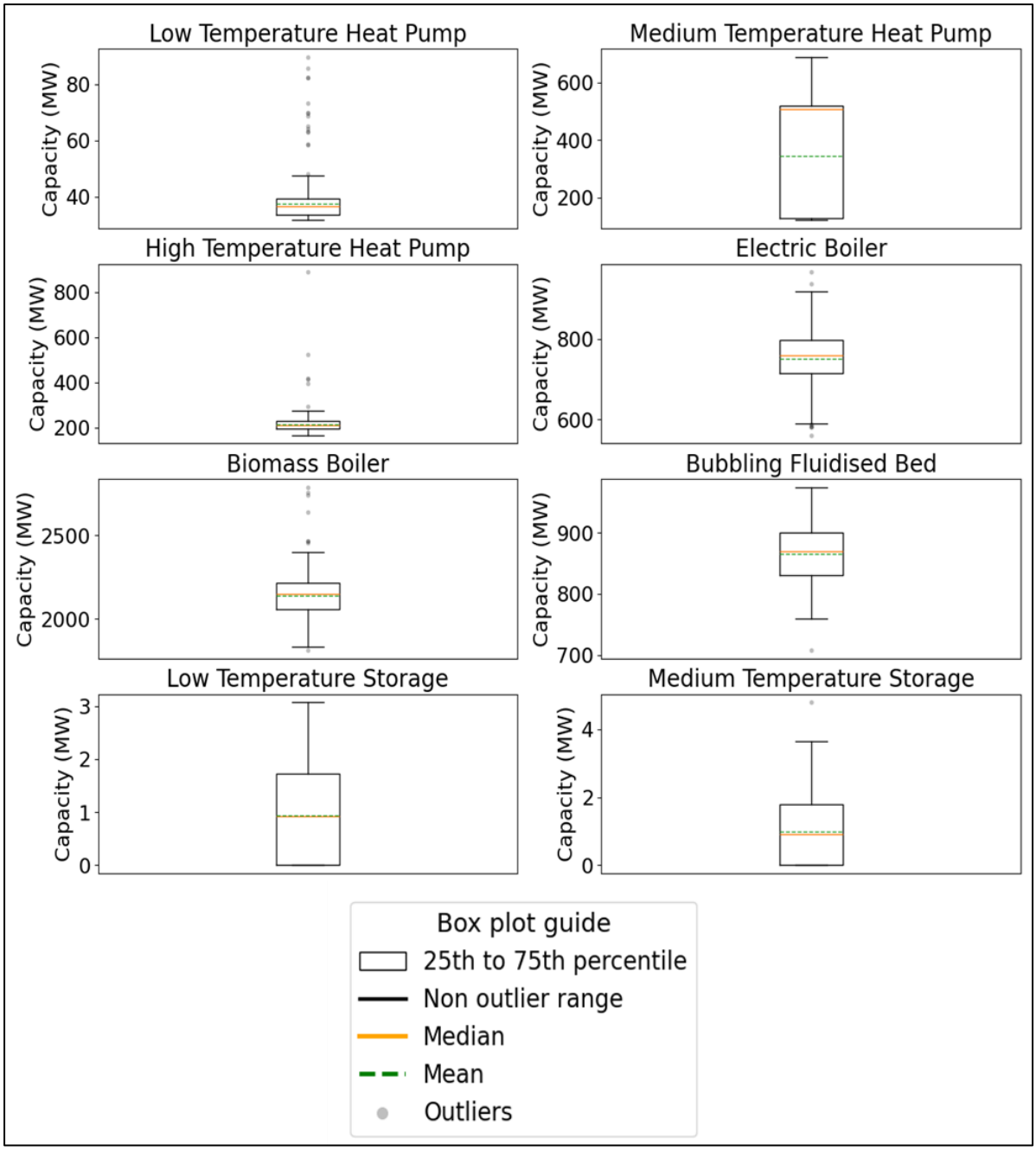


Figure 9.7: System-level results on heating technologies and thermal storage options.

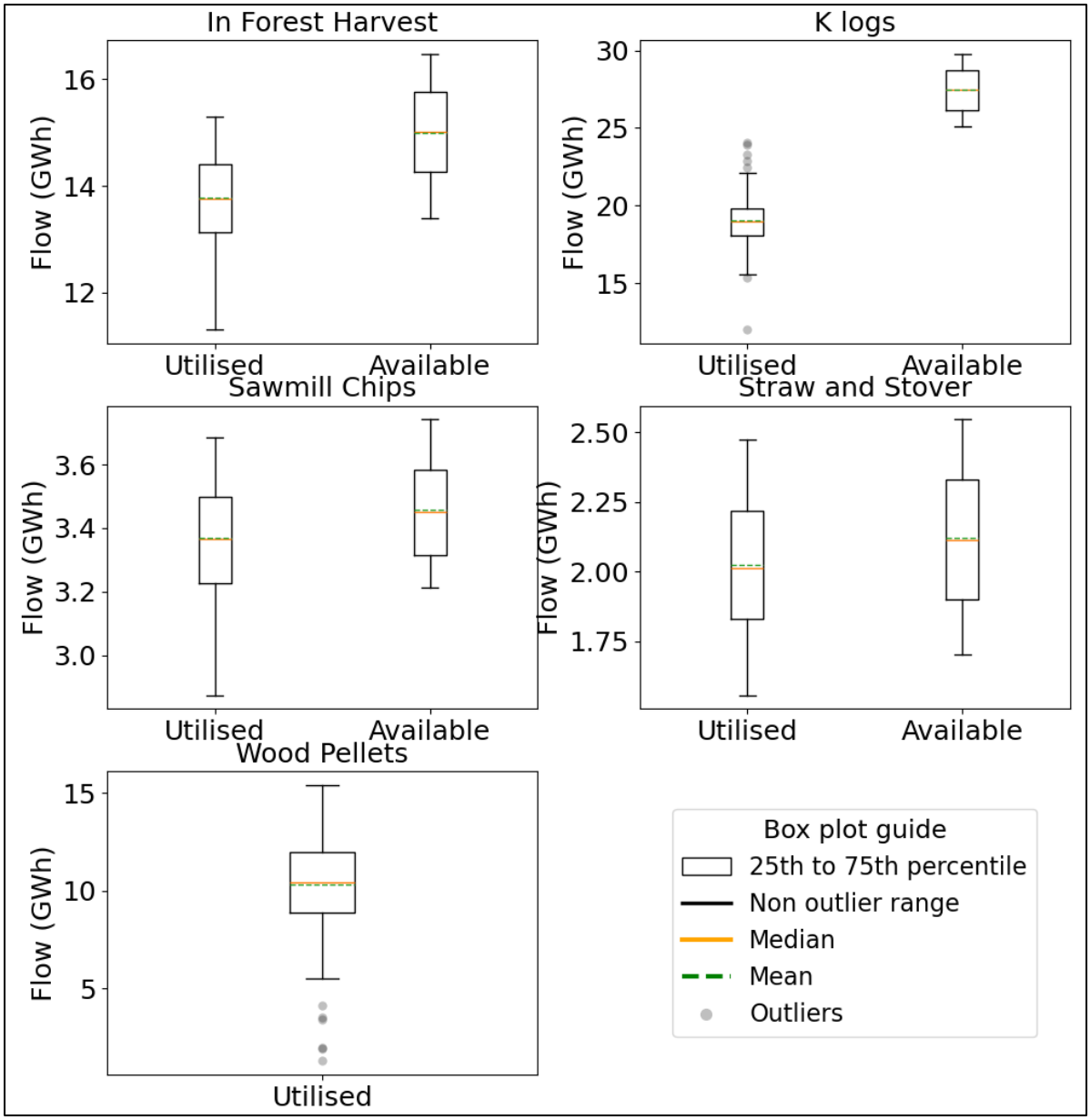


Figure 9.8: System-level results on biomass resource utilisation and availability.

Individual System View

As previously discussed, the modelling framework also allows a specified individual component with an assigned identifier within each layer to be extracted and analysed at specific locations due to the high spatial granularity of the model. This functionality enables detailed examination of technology deployment and fuel-switching behaviour at the factory level, which is particularly useful for energy modellers to support informed policymaking, as well as for industrial sites to identify the most suitable technologies for heat decarbonisation in most circumstances.

In this section, an example case study is presented by selecting Industrial ID 8 and Biomass Node ID 2 from the results database. The selected site corresponds to the Whareroa Fonterra dairy site (ID 8). This site is used as an example to analyse the fuel-switching capability and technology configuration at the plant level under the Monte Carlo uncertainty analysis framework. The framework allows users to observe how local operating conditions and regional resource availability influence the choice of heating technologies by examining the distribution of technology deployment and fuel inputs within the model.

Biomass utilisation remains consistent across all simulation runs, as all available biomass resources are fully consumed in the factory model. The result indicates a saturated use of biomass resources within the studied biomass node. In Figure 9.9, it can be observed that the utilised biomass volume is equal to the available biomass volume in each run, based on the width of each box plot having the same value for both available and utilised biomass resources. As a result, the box plot primarily reflects the variation introduced by the random sampling of biomass availability in the Monte Carlo uncertainty analysis.

Referring to Figure 9.10, LTHP and MTHP remained consistently fixed with a mean capacity of 3.4 MW and 11.64 respectively. This shows that the factory site is not affected by random sampling of electricity prices nor biomass specifications. There is a fuel-switching probability on HTHP and BB, where both could be potentially deployed based on different stochastic conditions. This is because there is high variability in the capacity installation for BB and HTHP from Figure 9.10.

Table 9.2: Factory specifications connected by biomass node ID=2.

Factory ID	Sector	Process Type	Distance (km)
9	Dairy	Dairy Processing (Milk Powder, Casein, Cheese)	41.9
23	Wood Processing	Wood Processing (MDF Panels)	50.3
29	Wood Processing	Timber Processing (Sawmill, Kilns)	35.6
55	Wood Processing	Timber Processing (Sawmill)	24.5
77	Dairy	Dairy Processing (Sheep Milk Powder, Infant Formula)	5.7
78	Meat	Meat Processing (Beef, Lamb)	12.5
96	Dairy	Dairy Processing (Cheese, Milk Powder)	13.8
130	Wood Processing	Meat Processing	30.0
178	Other	Textile Services (Linen Rental, Laundry)	9.4
195	Wood Processing	Wood Processing	5.5
237	Meat	Meat Processing (Beef, Lamb)	13.3
325	Other	Construction	29.2

The energy modelling framework is also capable of extracting results from the biomass perspective. From the results for biomass ID 2 to factory locations, it is observed that it provides solid biomass fuel resources to 12 major locations for heating. For Factory ID 9 and Factory ID 23, in forest harvest residues supply a mean of 1.95 MWh and 31. MWh respectively across the 24-hour horizon. Biomass supply is also detected at Factory ID 29 and Factory ID 55. Sawmill chips and K logs are present at Factory ID 77, although from Figure 9.11, sawmill chips are highly inconsistent based on the distribution statistics. For Factory 195, K logs would have a mean supply of 20.4 MWh, whereas sawmill chips show uncertainty with an interquartile range of 0 to 15.82 MWh. In the case of Factory ID 325, biomass resource flow is relatively high, with a total mean of all biomass resources of 57.3 MWh, which is the highest among all sites.

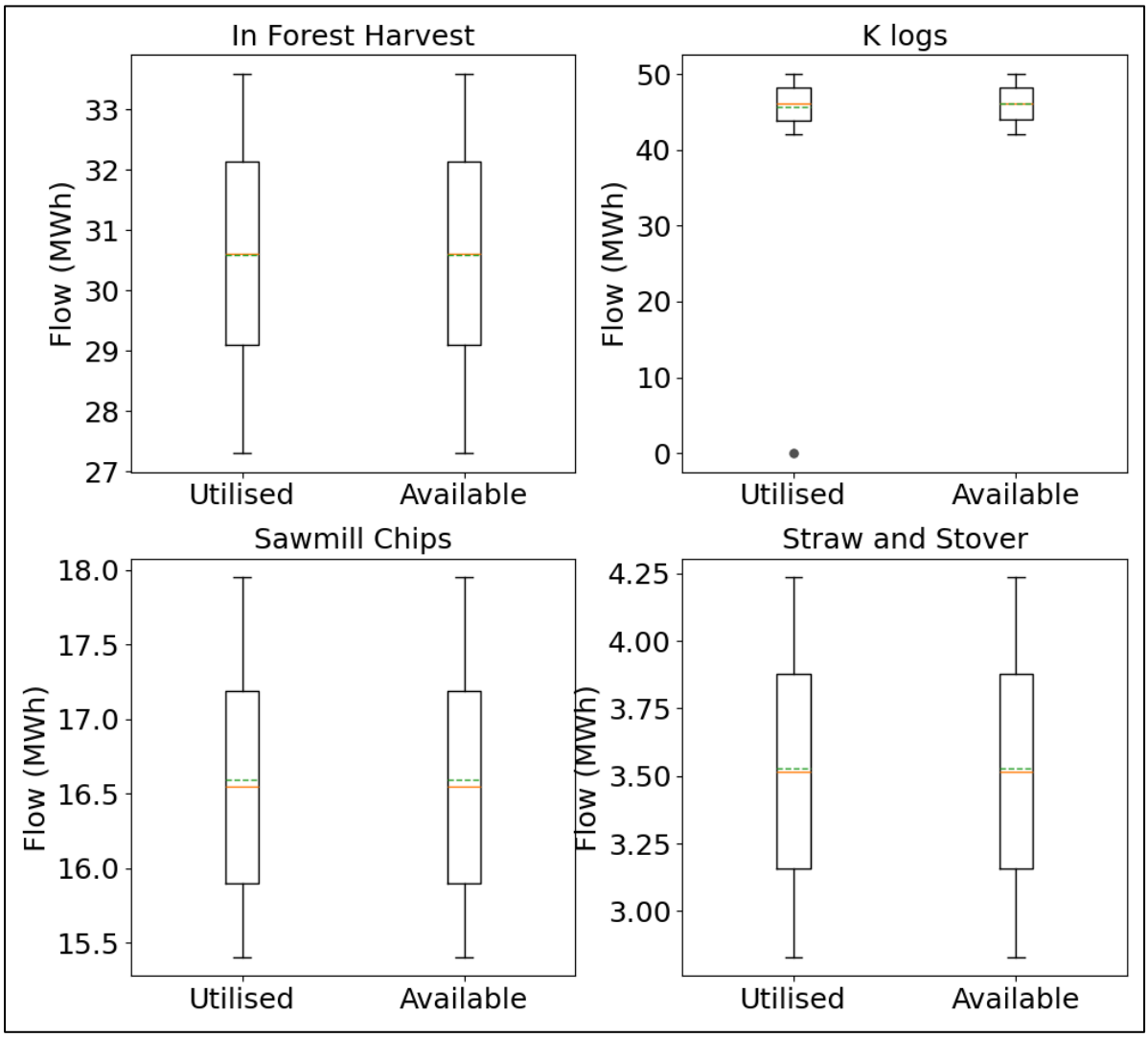


Figure 9.9: Individual biomass nodal ID=2 level results on biomass resource utilisation and availability.

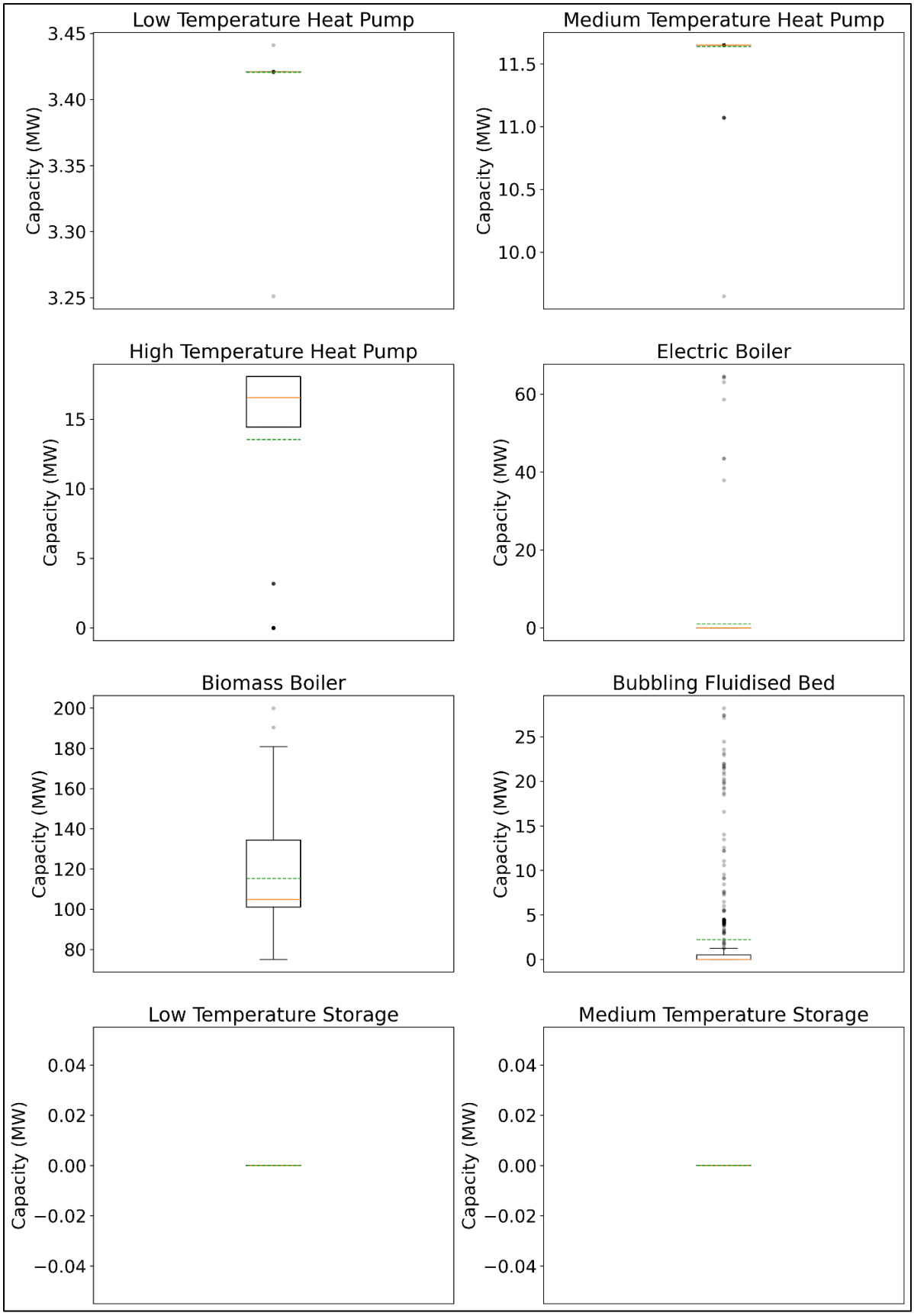


Figure 9.10: Individual industrial site ID=8 level perspective on heating technologies and thermal storage potential.

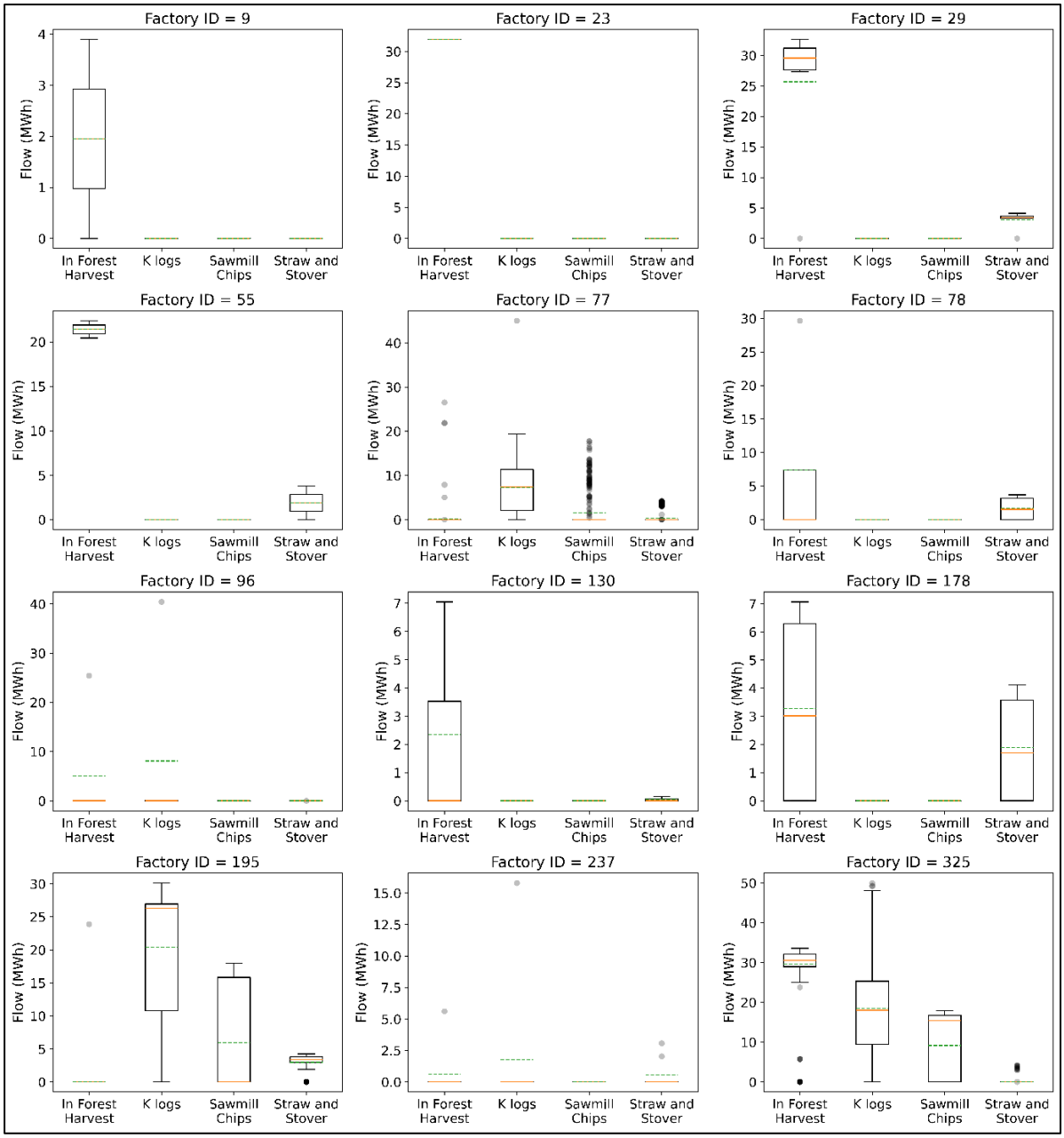


Figure 9.11: Individual biomass nodal ID=2 level results connection to 12 factory sites.

9.6 Chapter Summary

This chapter presented a method for dealing with uncertain parameters that can be applied in a highly spatial energy model. The random sampling of input parameters is conducted and analysed using iterative runs that are built within the P-graph representation framework and solved using the Pyomo modelling framework. Several insights can be evaluated from the modelling outcomes using the Monte Carlo method, particularly system-level trends in fuel switching for process heat technologies at major industrial sites, while also analysing how biomass resource availability and price influence the utilisation of biomass as a solid fuel in process heat decarbonisation.

From a high-level perspective, despite having a small share of the total capacity builds, LTHP and MTHP have shown reliable technology deployment outcomes with low variation, which cements the role of heat pumps in industrial sites. On the other hand, electric-based technologies such as MTHP and EB have high variability, as both are competing for their fuel-switching shares with BB and BFB. All biomass resource types are included in the model, regardless of the availability and type of biomass present. All biomass resources have a high utilisation rate, indicating that biomass plays an important role in the process heat in industry.

Looking at the site-level perspective, one biomass site and factory site are selected as an example. For the case of a biomass site, all available biomass resources are fully utilised in the model. This is because all factory locations are located considerably close to the biomass location, leading to logistical convenience when transporting biomass resources. This chapter presents several biomass allocation patterns across resources and factory sites.

This thesis evidences the influence of temporal variation at the annual-to-monthly resolution, demonstrated in Chapters 7 and 8. The extension of this comparison to sub-daily resolution was demonstrated at a representative scale in Chapter 9, using a 4-hourly Monte Carlo analysis over a representative daily horizon; extending this comparison to the full national model at sub-daily resolution remains constrained by computational tractability and is identified as future work.

Chapter 10: Conclusions and Recommendations for Future Work

Beyond its application to industrial decarbonisation, this thesis has introduced a methodological contribution to process systems engineering (PSE), which is a systematic approach for coupling graph-based process synthesis representations with algebraic disjunctive programming formulations. PNS methodologies such as P-graph offer structural clarity and combinatorial rigour but remain constrained in scalability due to their built-in branch-and-bound solvers, whereas algebraic optimisation environments offer powerful solvers but lack an intuitive representation and are prone to obscuring the link between formulation and solution, complicating debugging. Friedler et al. (2019) identified P-graph's features in sizing capacity and representing cascading effects to recognise alternative configurations as one of the field's key unresolved research challenges. Friedler et al. (2019) notes that this capability remains absent even from commercial flowsheet software. The P-graph-to-Pyomo porting tool developed in the thesis addresses this gap directly by providing a generalisable translation methodology that can be used in any P-graph problem, not only P-graph-generated energy modelling instances. This contribution of preserving the interpretability of the PNS representation while inheriting the scalability of a disjunctive programming formulation is transferable beyond the energy systems domain to any process synthesis problem for which PNS-based structural clarity and large-scale solver performance are both required.

This thesis developed a spatial energy modelling framework that integrates P-graph superstructure generation with automated translation into a Pyomo.GDP optimisation model. The framework enables structural optimisation of large-scale energy systems and supports multiple simulation modes for energy transition analysis.

The thesis set out to address a critical question in industrial decarbonisation: what fuel-switching strategy should each industrial site adopt in order to minimise cost from a system-wide perspective? The developed energy modelling framework is intended to support energy modellers, policymakers, and multi-site industrial companies in evaluating coordinated decarbonisation strategies. The framework explicitly accounts for biomass transport connections, regional supply constraints, electricity network interactions, and technology selection across multiple system layers. A follow-up research question on how electricity configurations change after accounting for non-industrial electricity demand and increased electrification in industrial process heat applications is addressed using a system-level grid reinforcement model that is interconnected with the factory model.

A key methodological contribution of this thesis is the integration of P-graph representation with a Pyomo-based optimisation environment. Traditional P-graph implementations are limited by built-in alternative MILP solvers such as CBC and GLPK, which restrict scalability for large, highly spatial problems. To overcome this limitation, a porting tool was developed to transform P-graph nodes and edges into a generalised disjunctive programming formulation within Pyomo. This

enables the use of advanced commercial solvers, and the disjunctive nature significantly improves computational performance for large-scale problems.

10.1 Section A: Energy Modelling Framework Creation

One of the aims of this research is the development of a harmonised modelling workflow that retains the visual clarity of P-graph while leveraging the flexibility and power of modern optimisation solvers. The porting tool converts the graphical superstructure representation into algebraic form, enabling disjunctive programming to be solved using MILP reformulation techniques such as Big-M transformation or convex hull formulation.

Disjunctive programming allows complex logical constraints and structural decisions to be represented explicitly, rather than relying solely on algebraic binary approximations within standard P-graph MILP formulations. This significantly enhances modelling flexibility when representing technology selection, mutual exclusivity, and directional flow constraints. In order to improve usability and flexibility, multiple optimisation modes were implemented within the integrated framework, including:

- Deterministic single-stage optimisation
- Deterministic sequential multi-stage optimisation
- Monte Carlo method

Moreover, solutions generated in Pyomo can be re-integrated into the P-graph environment if desired, preserving structural interpretability. The approach of using P-graph as building and visualisation of the model and porting the P-graph problem to be run in Pyomo solves both problems, which are:

- The difficulty of scaling large and complex problems directly within standalone P-graph solvers due to large combinatorial search space.
- The lack of structural visualisation and infeasibility diagnostics in algebraic modelling environments such as Pyomo.

By integrating both graphical superstructure modelling with algebraic optimisation, the framework achieves both interpretability and scalability.

Beyond the core integration of P-graph and Pyomo, the framework is supported by geospatial processing and output reporting modules which reinforces the practicality and usability of the case study setup while supporting clearer presentation of results.

10.2 Section B: Case Study Workflow and Comparative Analysis of Simulation Modes

As the energy modelling framework is developed using the generic P-graph method and Pyomo environment, the framework is inherently generalisable and can be applied to any case study. The generalisability operates at the level of the modelling architecture rather than the case study itself. The P-graph-to-Pyomo porting tool, the modular structures of components, the GDP formulation,

and the simulation modes are not tied to the resource base. Case studies can be modified by changing the entries in the input data, such as technology costs, resource availability, facility locations, spatial distances and grid constraints. The distinction matters because New Zealand has a comparatively favourable starting position for industrial decarbonisation studies, as the grid is already dominated by renewable hydroelectricity generation as discussed in Chapter 6 and simpler to define boundary conditions as an island nation.

Applying the framework to a region with a different energy landscape with higher fossil-fuel dominated grid, lower biomass availability would likely produce a different least-cost pathway, placing a greater expansion on electricity generation than the biomass-led solution found here. This would serve as a valuable stress test of both computational scalability and the framework's ability to surface region-appropriate trade-offs. The transferability speaks to the broader challenge of industrial decarbonisation, which admits that not a single global solution for fuel switching as the resource availability and infrastructure allowance vary between regions. Therefore, the contribution of this thesis therefore lies as much in providing a replicable, re-parameterisable methodology while returning specific New Zealand-based results.

Nevertheless, the case study of New Zealand, which is presented in the thesis, the solution provides several key insights. Results from the single deterministic multiperiod and sequential optimisation analyses indicate that biomass plays a dominant role in least-cost decarbonisation pathways, accounting for approximately 85–90% of decarbonised process heat under the baseline and sensitivity scenarios. Electrification contributes the remaining share, leading to an increase in industrial electricity end-use demand of approximately 3.5–9.2 TWh, depending on different scenarios such as changes in electricity sales price, biomass availability, and biomass price.

Biomass consumption for industrial heat ranges from approximately 7.7 to 13.5 TWh, significantly exceeding electrified heat demand. When combined with projected growth in residential, commercial, and transport electrification, total electricity generation increases from approximately 38.9 TWh in 2022 to approximately 46.2 TWh by 2037 under the base model scenario, which is an increase of approximately 7.7 TWh.

As the work is extended to grid reinforcement modelling, infrastructure upgrade costs are internalised through an adjusted electricity price coefficient. Under a Class 5 cost estimate framework, electricity prices are projected to change within a range of -12.6 to -2%, assuming uniform cost allocation across sectors. Comparative checks are conducted against known transmission upgrades and generation developments post-2022, providing qualitative validation of the model's reinforcement results.

Across repeated simulation runs, the model estimates the probability of electrification uptake at each industrial site and across sectors. The Monte Carlo results provide a distributional view of fuel-switching outcomes rather than a single deterministic solution, allowing the model to remain robust under uncertainty. The implementation of Monte Carlo analysis highlights electrification shares, biomass utilisation, and regional and spatial patterns under different input assumptions, which is useful for informing decisions on fuel-switching project deployment.

Taken together, the three simulation modes establish a claim that none could support alone. The deterministic multiperiod analysis shows that a biomass-dominant, temporally varying fuel-switching pathway is least-cost under a single, internally consistent set of assumptions. The sequential analysis shows that the pathway remains close to cost-optimal once the electricity system consequences of the electrification component are endogenized, with price adjustment of -12.6% to -2% under Class 5 cost estimate. The Monte Carlo analysis in Chapter 9 shows that the fuel-switching outcomes remain dominant configuration across a distribution of plausible electricity price, biomass availability, and biomass cost scenarios. Individually, each mode addresses a specific limitation of the others. Jointly, the three modes provide a level of assurance in the identified decarbonisation pathway, that it is cost-optimal, internally consistent with its own infrastructure requirements, as well as robust to key sources of uncertainty, and that no single simulation mode could establish in isolation.

Against the research gaps defined in Chapter 2, no reviewed tool combined spatial and structural detail with solver scalability. The results above show that the P-graph-to-Pyomo coupling achieves this combination for a national-scale case study, which is consistent with the challenge identified by Friedler et al. (2019). This was enabled by simplifications, specifically by a uniform grid cost coefficient and monthly temporal resolution. Across all three simulation modes, biomass consistently dominates in the process heat decarbonisation pathway, indicating this result reflects New Zealand's resource endowment rather than any single set of modelling assumptions. The methodological contribution is expected to transfer to any process synthesis problem requiring both structural clarity and solver scalability. The empirical findings, however, are more case study based. They depend on New Zealand's renewable-dominated grid, as discussed above. They are also capped at 500°C, which is the boundary for indirect process heating technologies.

10.3 Future Work on Energy Modelling Framework

The modelling framework developed in this thesis provides a comprehensive system-level perspective on industrial fuel-switching. However, several areas for future improvement can be identified.

First, the model does not conduct highly detailed and sophisticated structural system expansion decisions. In particular, the addition of new POCs, new substations, or load shedding mechanisms are not available in this current model. The framework evaluates reinforcement of existing infrastructure but does not endogenously optimise alternative network topologies. Extending the model to allow structural network expansion decisions would be helpful for transmission and distribution grid planning. This includes electricity firming where electricity is supplied across the grid using dispatchable generations when intermittent renewables are unable to meet demand.

In addition, the grid reinforcement representation focuses primarily on dominant equipment and proxy upgrades rather than a full engineering replication of the power system. Although this simplification improves computational tractability, future work could incorporate more detailed modelling of substation components, switchgear constraints, and bus-level power flow approximations. This would provide a more granular assessment of spatially differentiated

upgrade costs and allow more accurate attribution of reinforcement expenses at the distribution level.

As mentioned in Section 8.7, the current framework adjusts electricity prices using an averaged electricity price coefficient. However, future work could investigate how spatially differentiated upgrade costs should be allocated, particularly at the distribution or regional network level. A refined treatment could distinguish between local distribution upgrades and regional transmission reinforcements, enabling a clearer assessment of who bears the cost of electrification-driven investments.

The research question in this thesis focuses on overall system cost minimisation for fuel-switching. Therefore, the optimisation model returns solutions that are optimal from a system perspective rather than from the viewpoint of individual industrial sites. In practice, site-level decisions may differ due to financing constraints, contract type, or local electricity contracts. Future work could be extended using game-theoretic formulations that represent the decentralisation. A Stackelberg or bi-level formulation would be well suited for the scope, where a system-level player, such as the regulator which plays as a leader, sets a policy signal in a form of carbon price, subsidy or electricity tariff structure, while industrial sites respond as followers by independently minimising their own cost given that signal. Moreover, a Price-of Anarchy-type metric that compares the cost of the decentralised equilibrium to the system-optimal solution can be applied at the same time to industrial sites who act based on self-interests instead of considering the overall efficiency of the system.

Apart from fuel-switching decisions, the current formulation of shared infrastructure and resource costs also assumes a centrally coordinated allocation. In Chapter 8, grid reinforcement costs are assumed to be applied through a uniform, averaged electricity price adjustment applied across all sites. At the same time, multiple factories draw on the same limited biomass resources within a given region. Cooperative game theory, using solution concepts such as the Shapley value, could be introduced to determine a more equitable allocation of shared grid reinforcement costs among the sites that benefit from the given upgrade, rather than applying a uniform coefficient across the country. Similarly, competition for constrained biomass resources between neighbouring sites could be represented as a congestion or common-pool resource game, where sites decide independently how much biomass is drawn from a limited source. Under decentralised and uncoordinated demand, this could result in outcomes where nearby biomass sources over-subscribed relative to the system-optimal allocation identified in the thesis, where the outcome shows that biomass is dominant in decarbonising industrial process heat. Exploring these game-theoretic extensions could help clarify whether the thesis remain realistic, based on what coordination mechanisms, whether price-based, contractual, or allocation-based, would be needed to narrow the gap.

As machine learning and artificial intelligence have been deployed in large scale mathematical formulation with many successes, another gap that can be filled in this work to leverage the use of machine learning techniques such as supervised learning and unsupervised learning (Lin *et al.*, 2025). Supervised learning such as Support Vector Machines, Decision Trees, Random Forests and

k-Nearest Neighbours are widely used in energy modelling applications. It is often used in generating synthetic high-quality data which can be difficult to obtain (Maleki *et al.*, 2024). On the other hand, unsupervised learning algorithms are used to study weather patterns and capture different data trends which ultimately improve the forecasting of solar and wind energy (Salehmin *et al.*, 2024).

A future work closer to the scope would be to train a machine-learned surrogate model, such as a neural network or gradient-boosted model on a library of previously solved model instances to approximate near-optimal solutions or generate good initialisations for the GDP solver. This strategy would reduce the solve times for large scale or repeated runs, which would be relevant for the Monte Carlo simulation mode explained in Chapter 9, where the same model structure being solved across multiple iterations. Another structurally tailored approach would be to apply Graph Neural Networks, given that the P-graph superstructure is a native graph representation of materials, resource nodes and operating units. The neural networks can be trained to identify which regions within the superstructure are likely to generate a near-optimal solution, effectively pruning the combinatorial search space before it is passed to the solver. This ultimately reduces the number of subproblems which gives way of potentially further increasing the problem size. In both cases, such surrogate approaches would be intended to complement rather than actually replace the underlying optimisation formulation.

A future extension could develop a site-level optimisation model with a different objective function, allowing comparison between system-optimal and site-optimal decisions. Such comparative modelling could provide valuable insights to cover the gaps between private and high-level public optimisation outcomes.

Additional simulation modes can also be incorporated. For example, a Markov chain, which is a stochastic process, could be used to analyse sequences of possible future states and address uncertainties such as dry years, low biomass availability, and other impacts on model parameters. This stochastic optimisation would be inherently valuable in energy system modelling to test the reliability and robustness of the modelling results.

Additionally, the incorporation of risk management considerations could be strengthened. Although uncertainty is addressed through Monte Carlo analysis, future work could integrate robust optimisation approaches to explicitly account for downside risk in fuel prices, biomass availability, or operational handling on electricity network flow.

10.4 Future Case Study Extensions and Energy Modelling Architecture

From a technical implementation perspective, further enhancements could improve the robustness of the modelling backend. At present, the preprocessing steps are optional and remain exogenous, as input data may be drawn from different databases across case studies, each with varying structures and attributes. Model inputs are therefore prepared externally in Excel spreadsheets and manually transferred into the core energy modelling framework. Although this approach is functional, it could be improved through the development of an integrated preprocessing pipeline that standardises exogenous inputs, validates data formats, and directly

links input models across different case studies. Such a structured data pipeline would enhance reproducibility, reduce manual intervention, and improve overall workflow efficiency.

Although Renewables.ninja and ChromeDriver are open-source, the main geospatial tool currently used, ArcPy, requires a licence. For modellers without access to this licence, open-source alternatives such as QGIS and GeoPandas could be adopted. The development of geospatial preprocessing tools based on these libraries would make their integration into the core optimisation framework and data pipeline more justifiable, as the workflow could be streamlined without the risk of being constrained by proprietary software requirements. By integrating the preprocessing tool with the already integrated optimisation and output modules, manual labour would be reduced, the risk of human error lowered, and reproducibility improved. This initiative is particularly beneficial for users with limited coding experience or those working in a low-code environment.The graphical P-graph representation currently assumes linear relationships between nodes connected by edges. Future work could explore embedding non-linear representations within nodes, allowing certain processes to follow non-linear cost curves or performance relationships while maintaining overall model tractability.

Currently, modular components are built within the P-graph framework and outputs related to the components are extracted using masking techniques after passing through the solver for reporting in tables or diagrams. However, as modelling scope may vary between users, the framework should, in theory, allow for customisation through the generic P-graph to Pyomo porting tool. For example, a user could incorporate an additional component such as data centres, such as the ones potentially being built in Invercargill, which include both heating and cooling requirements, into the model. This flexibility highlights the adaptability of the framework. Nevertheless, the development of a more user-friendly front-end interface or coded modular components would further enhance usability, enabling users to more easily customise and implement technology blocks according to their specific modelling needs.

References

Achterberg, T. and Wunderling, R. (2013) “Mixed integer programming: Analyzing 12 years of progress,” in M. Jünger and G. Reinelt (eds.) *Facets of Combinatorial Optimization: Festschrift for Martin Grötschel*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 449–481. Available at: https://doi.org/10.1007/978-3-642-38189-8_18.

Akashi, O. *et al.* (2012) “GHG emission scenarios in Asia and the world: The key technologies for significant reduction,” *Energy Economics*, 34, pp. S346–S358. Available at: https://doi.org/10.1016/j.eneco.2012.04.011.

Al-Hanoot, A.K. *et al.* (2025) “Monte Carlo simulation for real-world energy yield analysis of car park solar PV system installations in harsh environments,” *Results in Engineering*, 28, p. 106996. Available at: https://doi.org/10.1016/j.rineng.2025.106996.

Al-Wakeel, A., Wu, J. and Jenkins, N. (2017) “k-means based load estimation of domestic smart meter measurements,” *Applied Energy*, 194, pp. 333–342. Available at: https://doi.org/10.1016/j.apenergy.2016.06.046.

Arpagaus, C. *et al.* (2018) “High temperature heat pumps: Market overview, state of the art, research status, refrigerants, and application potentials,” *Energy*, 152, pp. 985–1010. Available at: https://doi.org/10.1016/j.energy.2018.03.166.

Aryanpur, V. *et al.* (2021) “A review of spatial resolution and regionalisation in national-scale energy systems optimisation models,” *Energy Strategy Reviews*, 37, p. 100702. Available at: https://doi.org/10.1016/j.esr.2021.100702.

Ata, M.Y. (2007) “A convergence criterion for the Monte Carlo estimates,” *Simulation Modelling Practice and Theory*, 15(3), pp. 237–246. Available at: https://doi.org/10.1016/j.simpat.2006.12.002.

Augutis, J. *et al.* (2020) “Analysis of energy security level in the Baltic States based on indicator approach,” *Energy*, 199, p. 117427. Available at: https://doi.org/10.1016/j.energy.2020.117427.

Aurecon (2024) *Energy technology cost and technical parameter review report*. Technical Report. Aurecon.

Aviso, K.B. *et al.* (2017) “A P-graph model for multi-period optimization of sustainable energy systems,” *Journal of Cleaner Production*, 161, pp. 1338–1351. Available at: https://doi.org/10.1016/j.jclepro.2017.06.044.

Badakhsh, A. and Mothilal Bhagavathy, S. (2024) “Caveats of green hydrogen for decarbonisation of heating in buildings,” *Applied Energy*, 353, p. 122152. Available at: https://doi.org/10.1016/j.apenergy.2023.122152.

Bahl, B. *et al.* (2018) “Typical Periods for Two-Stage Synthesis by Time-Series Aggregation with Bounded Error in Objective Function,” *Frontiers in Energy Research*, 5(35). Available at: https://doi.org/10.3389/fenrg.2017.00035.

Balas, E. (2018) *Disjunctive Programming*. 1st ed. Cham: Springer. Available at: https://doi.org/10.1007/978-3-030-00148-3.

Bale, C.S.E., Varga, L. and Foxon, T.J. (2015) “Energy and complexity: New ways forward,” *Applied Energy*, 138, pp. 150–159. Available at: https://doi.org/10.1016/j.apenergy.2014.10.057.

Balyk, O. *et al.* (2022) “TIM: modelling pathways to meet Ireland’s long-term energy system challenges with the TIMES-Ireland Model (v1.0),” *Geoscientific Model Development*, 15(12), pp. 4991–5019. Available at: https://doi.org/10.5194/gmd-15-4991-2022.

Barasa, M. *et al.* (2018) “A cost optimal resolution for Sub-Saharan Africa powered by 100% renewables in 2030,” *Renewable and Sustainable Energy Reviews*, 92, pp. 440–457. Available at: https://doi.org/10.1016/j.rser.2018.04.110.

Barbhuiya, S., Bhusan Das, B. and Adak, D. (2024) “Roadmap to a net-zero carbon cement sector: Strategies, innovations and policy imperatives,” *Journal of Environmental Management*, 359, p. 121052. Available at: https://doi.org/10.1016/j.jenvman.2024.121052.

Bataille, C. *et al.* (2018) “A review of technology and policy deep decarbonization pathway options for making energy-intensive industry production consistent with the Paris Agreement,” *Journal of Cleaner Production*, 187, pp. 960–973. Available at: https://doi.org/10.1016/j.jclepro.2018.03.107.

Beale, E.M.L. (1979) “Branch and Bound Methods for Mathematical Programming Systems,” in P.L. Hammer, E.L. Johnson, and B.H. Korte (eds.) *Annals of Discrete Mathematics*. Elsevier, pp. 201–219. Available at: https://doi.org/10.1016/S0167-5060(08)70351-0.

Beccarello, M. and Di Foggia, G. (2023) “Review and Perspectives of Key Decarbonization Drivers to 2030,” *Energies*, 16(3). Available at: https://doi.org/10.3390/en16031345.

Beck, T. *et al.* (2016) “Assessing the influence of the temporal resolution of electrical load and PV generation profiles on self-consumption and sizing of PV-battery systems,” *Applied Energy*, 173, pp. 331–342. Available at: https://doi.org/10.1016/j.apenergy.2016.04.050.

Begemann, T. *et al.* (2021) *Industriewärme klimaneutral: Strategien und Voraussetzungen für die Transformation*. Report RWTH-2021-06369. RWTH Publications.

Belotti, P. (2013) “Bound reduction using pairs of linear inequalities,” *Journal of Global Optimization*, 56(3), pp. 787–819. Available at: https://doi.org/10.1007/s10898-012-9848-9.

Belotti, P. *et al.* (2013) “Mixed-integer nonlinear optimization,” *Acta Numerica*, 22. Available at: https://doi.org/10.1017/S0962492913000032.

Bernal, D.E. *et al.* (2022) “Advances in Generalized Disjunctive and Mixed-Integer Nonlinear Programming Algorithms and Software for Superstructure Optimization,” in Y. Yamashita and M. Kano (eds.) *Computer Aided Chemical Engineering*. Elsevier, pp. 1285–1290. Available at: https://doi.org/10.1016/B978-0-323-85159-6.50214-1.

Bever, P. *et al.* (2024) “High-Temperature Heat Pumps for Industrial Use,” *Chemie Ingenieur Technik*, 96. Available at: https://doi.org/10.1002/cite.202300241.

Bioenergy Association of New Zealand (2011) *Heat Plant Database*. Available at: https://www.bioenergy.org.nz/resource/heat-plant-database (Accessed: February 25, 2026).

Bioenergy Association of New Zealand (2021) *Pricing of Different Biomass Fuels*. Bioenergy Association. Available at: https://www.bioenergy.org.nz/documents/admin/TNSB74-Pricing-of-different-biomass-fuels.pdf.

Bless, F. *et al.* (2017) “Theoretical analysis of steam generation methods - Energy, CO2 emission, and cost analysis,” *Energy*, 129, pp. 114–121. Available at: https://doi.org/10.1016/j.energy.2017.04.088.

Bogdanov, D. and Solomon, O. (2023) “Hierarchical approach to energy system modelling: Complexity reduction with minor changes in results,” *Energy*, 273, p. 127213. Available at: https://doi.org/10.1016/j.energy.2023.127213.

Boston Consulting Group (2025) *Energy to Grow: Securing New Zealand’s Future*. Boston Consulting Group. Available at: https://web-assets.bcg.com/27/08/7fcb68024dcc92f2ba01527e8623/energy-to-grow-full-report.pdf.

Bures, M. and Filipsky, M. (2016) “SmartDriver: Extension of Selenium WebDriver to Create More Efficient Automated Tests,” *2016 6th International Conference on IT Convergence and Security (ICITCS)*, pp. 1–4. Available at: https://doi.org/10.1109/ICITCS.2016.7740370.

Butera, G. *et al.* (2020) “Thermodynamic analysis of methanol synthesis combining straw gasification and electrolysis via the low temperature circulating fluid bed gasifier and a char bed gas cleaning unit,” *Energy*, 199, p. 117405. Available at: https://doi.org/10.1016/j.energy.2020.117405.

Cao, K.-K. *et al.* (2019) “Classification and Evaluation of Concepts for Improving the Performance of Applied Energy System Optimization Models,” *Energies*, 12(24). Available at: https://doi.org/10.3390/en12244656.

Cao, Z. *et al.* (2017) “Elaborating the History of Our Cementing Societies: An in-Use Stock Perspective,” *Environmental Science & Technology*, 51(19), pp. 11468–11475. Available at: https://doi.org/10.1021/acs.est.7b03077.

Carrion, M. and Arroyo, J. (2006) “A Computationally Efficient Mixed-Integer Linear Formulation for the Thermal Unit Commitment Problem,” *IEEE Transactions on Power Systems*, 21, pp. 1371–1378. Available at: https://doi.org/10.1109/TPWRS.2006.876672.

Chateau, B. and Lapillonne, B. (1990) “Accounting and end-use models,” *Energy*, 15(3), pp. 261–278. Available at: https://doi.org/10.1016/0360-5442(90)90088-J.

Chen, Q. *et al.* (2022) “Pyomo.GDP: an ecosystem for logic based modeling and optimization development,” *Optimization and Engineering*, 23, pp. 1–36. Available at: https://doi.org/10.1007/s11081-021-09601-7.

Clark, G. *et al.* (2024) “Assessment of fuel switching as a decarbonization strategy in the cement sector,” *Energy Conversion and Management*, 312, p. 118585. Available at: https://doi.org/10.1016/j.enconman.2024.118585.

Cleaver-Brooks (2021) *Electric & Electrode Boilers*. Available at: https://cleaverbrooks.com/Catalog/boilers/electric-and-electrode (Accessed: February 25, 2026).

Cococcioni, M. and Fiaschi, L. (2021) “The Big-M method with the numerical infinite M,” *Optimization Letters*, 15(7), pp. 2455–2468. Available at: https://doi.org/10.1007/s11590-020-01644-6.

Connolly, D. *et al.* (2010) “A review of computer tools for analysing the integration of renewable energy into various energy systems,” *Applied Energy*, 87, pp. 1059–1082. Available at: https://doi.org/10.1016/j.apenergy.2009.09.026.

Costa Dilmé, S. *et al.* (2024) “Multi-physics operation and sizing optimisation in Pyomo: Application to large irrigation systems.” *2024 Open Source Modelling and Simulation of Energy Systems (OSMSES)*, pp. 1–6. Available at: https://doi.org/10.1109/OSMSES62085.2024.10668997.

Dalena, F. *et al.* (2018) “Chapter 1 - Methanol Production and Applications: An Overview,” in A. Basile and F. Dalena (eds.) *Methanol*. Elsevier, pp. 3–28. Available at: https://doi.org/10.1016/B978-0-444-63903-5.00001-7.

Dananjaya, V. *et al.* (2024) “Synthesis, properties, applications, 3D printing and machine learning of graphene quantum dots in polymer nanocomposites,” *Progress in Materials Science*, 144, p. 101282. Available at: https://doi.org/10.1016/j.pmatsci.2024.101282.

Dancker, J. *et al.* (2019) “Optimal design and operation of a CHP based district heating system including a heat storage and electrode boiler to increase self-consumption,” *2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, pp. 1–6. Available at: https://doi.org/10.1109/ISGT-LA.2019.8895427.

Dantzig, G.B. (1951) “Maximization of a Linear Function of Variables Subject to Linear Inequalities,” in T.C. Koopmans (ed.) *Activity Analysis of Production and Allocation*. New York: Wiley and Chapman-Hall, pp. 339–347.

DeCarolis, J. *et al.* (2017) “Formalizing best practice for energy system optimization modelling,” *Applied Energy*, 194, pp. 184–198. Available at: https://doi.org/10.1016/j.apenergy.2017.03.001.

EECA (2019) *Electrode and electric resistance steam generators and hot water heaters for low-carbon process heating*. EEC0146. Energy Efficiency and Conservation Authority. Available at: https://www.genless.govt.nz/assets/Business-Resources/Electrode-electric-resistance-steam-generators-hot-water-heaters-for-low-carbon-process-heating.pdf.

EECA (2024) *Energy End Use Database*. Te Kāwanatanga o Aotearoa / New Zealand Government. Available at: https://www.eeca.govt.nz/insights/data-tools/energy-end-use-database/ (Accessed: February 25, 2026).

EECA (2025a) "How heat recovery is cutting costs and gas use in food manufacturing." Available at: https://www.eeca.govt.nz/insights/case-studies-and-articles/how-heat-recovery-is-cutting-costs-and-gas-use-in-food-manufacturing/.

EECA (2025b) *Regional Energy Transition Accelerator Dashboard*. Available at: https://www.eeca.govt.nz/insights/data-tools/regional-energy-transition-accelerator-dashboard/ (Accessed: February 25, 2026).

EECA (2025c) "Regional Energy Transition Accelerator (RETA)." Available at: https://www.eeca.govt.nz/co-funding-and-support/products/about-reta/ (Accessed: February 25, 2026).

Egerer, J. (2016) *Open Source Electricity Model for Germany (ELMOD-DE): Data documentation*. 83. DIW Berlin, p. 45. Available at: https://doi.org/10.18723/diw_ddc:2016-83.

Eichfelder, G. *et al.* (2021) "A general branch-and-bound framework for continuous global multiobjective optimization," *Journal of Global Optimization*, 80(1), pp. 195–227. Available at: https://doi.org/10.1007/s10898-020-00984-y.

Electricity Authority (2025) *EMI: Electricity Market Information*. Available at: https://www.emi.ea.govt.nz/ (Accessed: February 25, 2026).

Electricity Authority (2026) *New Zealand's electricity sector*. Available at: https://www.ea.govt.nz/your-power/new-zealands-electricity-sector/ (Accessed: February 25, 2026).

Electricity Offers NZ (2025) "Electricity Offers New Zealand: Compare Power Companies and Plans." Available at: https://electricityoffers.co.nz/.

Éles, A., Heckl, I. and Cabezas, H. (2024) "Modeling of a Biomass-Based Energy Production Case Study Using Flexible Inputs with the P-Graph Framework," *Energies*, 17(3). Available at: https://doi.org/10.3390/en17030687.

Energy Information Administration (2025) *Model documentation report: Industrial sector demand module of the National Energy Modeling System*. Washington, DC: U.S. Department of Energy, p. 235. Available at: https://www.eia.gov/outlooks/aeo/nems/documentation/industrial/pdf/IDM_AEO2025.pdf.

Ergo Consulting (2024) *Waikato Spare Capacity and Load Characteristics Report*. 23163-RPT-005. Prepared for Energy Efficiency and Conservation Authority (EECA).

Ergo Consulting (2025) *Manawatū-Whanganui Spare Capacity and Load Characteristics Report*. 23163-RPT-005. Prepared for Energy Efficiency and Conservation Authority (EECA).

European Commission (2008) *NACE Rev. 2: Statistical classification of economic activities in the European Community*. Luxembourg: Office for Official Publications of the European Communities (KS-RA-07-015-EN-N).

European Heating Industry (EHI) (2021) *Heating Market Report 2021*. European Heating Industry. Available at: https://ehi.eu/wp-content/uploads/2022/09/EHI-2021-Heating-Market-Report.pdf.

Fleiter, T., Worrell, E. and Eichhammer, W. (2011) "Barriers to energy efficiency in industrial bottom-up energy demand models—A review," *Renewable and Sustainable Energy Reviews*, 15(6), pp. 3099–3111. Available at: https://doi.org/https://doi.org/10.1016/j.rser.2011.03.025.

Fonterra NZ (2022) *140 years on and still going strong*. Available at: https://www.fonterra.com/nz/en/our-stories/articles/140-years-on-and-still-going-strong.html (Accessed: February 25, 2026).

Forschungsgesellschaft für Energiewirtschaft (FfE) (2021) "Normalized Industrial Electrical Load Profiles (Germany)." Munich, Germany: FfE München. Available at: https://opendata.ffe.de/dataset/normalized-industrial-electrical-load-profiles-germany/ (Accessed: February 25, 2026).

Franco, A. and Rocca, M. (2024) "Industrial Decarbonization through Blended Combustion of Natural Gas and Hydrogen," *Hydrogen*, 5(3), pp. 519–539. Available at: https://doi.org/10.3390/hydrogen5030029.

Fremstad, A., Underwood, A. and Zahran, S. (2018) "The Environmental Impact of Sharing: Household and Urban Economies in CO2 Emissions," *Ecological Economics*, 145, pp. 137–147. Available at: https://doi.org/10.1016/j.ecolecon.2017.08.024.

Frew, B.A. and Jacobson, M.Z. (2016) "Temporal and spatial tradeoffs in power system modeling with assumptions about storage: An application of the POWER model," *Energy*, 117, pp. 198–213. Available at: https://doi.org/10.1016/j.energy.2016.10.074.

Friedler, F. *et al.* (1992) "Graph-theoretic approach to process synthesis: axioms and theorems," *Chemical Engineering Science*, 47(8), pp. 1973–1988. Available at: https://doi.org/10.1016/0009-2509(92)80315-4.

Friedler, F. *et al.* (2019) "Prospects and challenges for chemical process synthesis with P-graph," *Current Opinion in Chemical Engineering*, 26, pp. 58–64. Available at: https://doi.org/https://doi.org/10.1016/j.coche.2019.08.007.

Friedler, F., Orosz, A. and Pimentel, J. (2022) *P-graphs for Process Systems Engineering: Mathematical Models and Algorithms*. 1st ed. Switzerland: Springer Cham. Available at: https://doi.org/10.1007/978-3-030-92216-0.

Frysztacki, M.M. *et al.* (2021) "The strong effect of network resolution on electricity system models with high shares of wind and solar," *Applied Energy*, 291. Available at: https://doi.org/10.1016/j.apenergy.2021.116726.

Frysztacki, M.M., Recht, G. and Brown, T. (2022) "A comparison of clustering methods for the spatial reduction of renewable electricity optimisation models of Europe," *Energy Informatics*, 5(1), p. 4. Available at: https://doi.org/10.1186/s42162-022-00187-7.

Fujimori, S. *et al.* (2019) "Energy transformation cost for the Japanese mid-century strategy," *Nature Communications*, 10(1). Available at: https://doi.org/10.1038/s41467-019-12730-4.

Fukunaga, A.S. (2011) "A branch-and-bound algorithm for hard multiple knapsack problems," *Annals of Operations Research*, 184(1), pp. 97–119. Available at: https://doi.org/10.1007/s10479-009-0660-y.

Gabrielli, P. *et al.* (2019) "Robust and optimal design of multi-energy systems with seasonal storage through uncertainty analysis," *Applied Energy*, 238, pp. 1192–1210. Available at: https://doi.org/10.1016/j.apenergy.2019.01.064.

García Vera, Y.E., Dufo-López, R. and Bernal-Agustín, J.L. (2019) "Energy Management in Microgrids with Renewable Energy Sources: A Literature Review," *Applied Sciences*, 9(18). Available at: https://doi.org/10.3390/app9183854.

García-Muñoz, F., Dávila, S. and Quezada, F. (2023) "A Benders decomposition approach for solving a two-stage local energy market problem under uncertainty," *Applied Energy*, 329, p. 120226. Available at: https://doi.org/10.1016/j.apenergy.2022.120226.

Gebremeskel, D., Ahlgren, E. and Bekele, G. (2023) "Long-term electricity supply modelling in the context of developing countries: The OSeMOSYS-LEAP soft-linking approach for Ethiopia," *Energy Strategy Reviews*, 45, p. 101045. Available at: https://doi.org/10.1016/j.esr.2022.101045.

Groissböck, M. (2019) "Are open source energy system optimization tools mature enough for serious use?," *Renewable and Sustainable Energy Reviews*, 102, pp. 234–248. Available at: https://doi.org/10.1016/j.rser.2018.11.020.

Groß, T. *et al.* (2023) "ETHOS.FINE: A Framework for Integrated Energy System Assessment." Available at: https://arxiv.org/abs/2311.05930 (Accessed: February 25, 2026).

Grossmann, I. (2021) *Advanced Optimization for Process Systems Engineering*. Available at: https://doi.org/10.1017/9781108917834.

Gungor, A. (2013) “Simulation of co-firing coal and biomass in circulating fluidized beds,” *Global Conference on Renewable energy and Energy Efficiency for Desert Regions 2011 “GCREEDER 2011,”* 65, pp. 574–579. Available at: https://doi.org/10.1016/j.enconman.2012.09.023.

Guo, J.-X. (2022) “Retrofitting strategy for biomass co-fired power plant,” *Clean Technologies and Environmental Policy*, 24(8), pp. 2531–2545. Available at: https://doi.org/10.1007/s10098-022-02332-y.

Gurobi Optimization, LLC (2026) “Gurobi Optimizer Reference Manual.” Available at: https://www.gurobi.com.

Hall, P. (2022) *Residual Biomass Fuel Projections for New Zealand; 2021: Indicative Availability by Region and Source*. PAD output number 45753044. Rotorua, New Zealand: Scion (New Zealand Forest Research Institute Limited). Available at: https://www.usewoodfuel.org.nz/documents/resource/Woody-biomass-residues-and-resources-2021-Feb2022_V5.pdf.

Hall, P. (2024) *Residual Biomass Fuel Projections for New Zealand; 2024: Indicative Availability by Region and Source*. 92781397. New Zealand: IEA Bioenergy. Available at: https://www.ieabioenergy.com/wp-content/uploads/2024/11/NZ-Woody-Biomass-Residues-and-Resources-2024.pdf.

Heckl, I. *et al.* (2015) “Process synthesis involving multi-period operations by the P-graph framework,” *24th European Symposium on Computer Aided Process Engineering(ESCAPE)*, 83, pp. 157–164. Available at: https://doi.org/10.1016/j.compchemeng.2015.04.037.

Hensen, M. (2025) *Level Playing Field measures: Electricity Authority Options Paper*. NZIER. Available at: https://www.ea.govt.nz/documents/7361/Major_Electricity_Users_Group_-_Report_uaGvxp0.pdf.

Heuberger, C.F. *et al.* (2017) “Power capacity expansion planning considering endogenous technology cost learning,” *Applied Energy*, 204, pp. 831–845. Available at: https://doi.org/10.1016/j.apenergy.2017.07.075.

Hilpert, S. *et al.* (2018) “The Open Energy Modelling Framework (oemof) - A new approach to facilitate open science in energy system modelling,” *Energy Strategy Reviews*, 22, pp. 16–25. Available at: https://doi.org/10.1016/j.esr.2018.07.001.

Hoffmann, M. *et al.* (2020) “A Review on Time Series Aggregation Methods for Energy System Models,” *Energies*, 13, p. 641. Available at: https://doi.org/10.3390/en13030641.

Hoffmann, M. *et al.* (2024) “A review of mixed-integer linear formulations for framework-based energy system models,” *Advances in Applied Energy*, 16. Available at: https://doi.org/10.1016/j.adapen.2024.100190.

Hohmann, M., Evins, R. and Lygeros, J. (2017) “Optimal dispatch of large multi-carrier energy networks considering energy conversion functions,” *Energy Procedia*, 122, pp. 80–85. Available at: https://doi.org/10.1016/j.egypro.2017.07.311.

HOMER Energy (2016) *HOMER Pro Version 3.7 User Manual*. Boulder, Colorado: HOMER Energy LLC. Available at: https://www.homerenergy.com.

Hörsch, J. *et al.* (2018) “PyPSA-Eur: An open optimisation model of the European transmission system,” *Energy Strategy Reviews*, 22, pp. 207–215. Available at: https://doi.org/10.1016/j.esr.2018.08.012.

Hörsch, J. and Brown, T. (2017) “The role of spatial scale in joint optimisations of generation and transmission for European highly renewable scenarios.” Available at: https://doi.org/10.48550/arXiv.1705.07617.

Howells, M. *et al.* (2011) “OSeMOSYS: The Open Source Energy Modeling System: An introduction to its ethos, structure and development,” *Energy Policy*, 39, pp. 5850–5870. Available at: https://doi.org/10.1016/j.enpol.2011.06.033.

Huang, H. *et al.* (2023) “Biomass briquette fuel, boiler types and pollutant emissions of industrial biomass boiler: A review,” *Particuology*, 77, pp. 79–90. Available at: https://doi.org/10.1016/j.partic.2022.08.016.

Huang, L. *et al.* (2021) *Branch and Bound in Mixed Integer Linear Programming Problems: A Survey of Techniques and Trends*. Available at: https://doi.org/10.48550/arXiv.2111.06257.

Huangfu, Q. and Hall, J.A.J. (2018) “Parallelizing the dual revised simplex method,” *Mathematical Programming Computation*, 10(1), pp. 119–142. Available at: https://doi.org/10.1007/s12532-017-0130-5.

Huuki, H. *et al.* (2022) “The economic cost of hydropower environmental constraints under decreasing price volatility,” *River Research and Applications*, 38(10), pp. 1815–1828. Available at: https://doi.org/10.1002/rra.4049.

IEA (2012) *Renewable Energy Technologies: Cost Analysis Series – Hydropower*. Cost Analysis Report. Abu Dhabi: International Renewable Energy Agency (IRENA).

IEA (2018) *Clean and efficient heat for industry*. Available at: https://www.iea.org/commentaries/clean-and-efficient-heat-for-industry.

IEA (2020) *Energy consumption in the iron and steel sector by scenario*. Available at: https://www.iea.org/data-and-statistics/charts/energy-consumption-in-the-iron-and-steel-sector-by-scenario (Accessed: February 25, 2026).

IEA (2024) *Renewables 2023*. Paris: IEA. Available at: https://www.iea.org/reports/renewables-2023.

IEA Bioenergy (2024) *Implementation of Bioenergy in the European Union – 2024 Update*. IEA Bioenergy. Available at: https://www.ieabioenergy.com/wp-content/uploads/2025/01/CountryReport2024_EU27_final_v2.pdf.

Ishak, S.A. and Hashim, H. (2015) "Low carbon measures for cement plant – a review," *Journal of Cleaner Production*, 103, pp. 260–274. Available at: https://doi.org/10.1016/j.jclepro.2014.11.003.

Isik, M., Dodder, R. and Kaplan, P.O. (2021) "Transportation emissions scenarios for New York City under different carbon intensities of electricity and electric vehicle adoption rates," *Nature Energy*, 6(1), pp. 92–104. Available at: https://doi.org/10.1038/s41560-020-00740-2.

Jani, D.B. *et al.* (2020) "A review on use of TRNSYS as simulation tool in performance prediction of desiccant cooling cycle," *Journal of Thermal Analysis and Calorimetry*, 140(5), pp. 2011–2031. Available at: https://doi.org/10.1007/s10973-019-08968-1.

Jenkin, T. *et al.* (2019) *Estimating the Impact of Residual Value for Electricity Generation Plants on Capital Recovery, Levelized Cost of Energy, and Cost to Consumers*. NREL/TP-6A20-72217. Golden, CO: National Renewable Energy Laboratory. Available at: https://www.nrel.gov/docs/fy19osti/72217.pdf.

Ji, M. *et al.* (2023) "Optimisation of multi-period renewable energy systems with hydrogen and battery energy storage: A P-graph approach," *Energy Conversion and Management*, 281, p. 116826. Available at: https://doi.org/10.1016/j.enconman.2023.116826.

Jochem, E. *et al.* (2024) "Energy Demand and Modelling of Energy Systems: Five Decades from Little Knowledge to Differentiated Know-How," in J. Edler and R. Walz (eds.) *Systems and Innovation Research in Transition: Research Questions and Trends in Historical Perspective*. Cham: Springer Nature Switzerland, pp. 237–273. Available at: https://doi.org/10.1007/978-3-031-66100-6_9.

Joyo, F.H., Nastasi, B. and Astiaso Garcia, D. (2025) "Decarbonization pathways for the pulp and paper industry: A comprehensive review," *Renewable and Sustainable Energy Reviews*, 223, p. 116070. Available at: https://doi.org/10.1016/j.rser.2025.116070.

Kannan, R. and Turton, H. (2012) "A Long-Term Electricity Dispatch Model with the TIMES Framework," *Environmental Modeling & Assessment*, 18, pp. 325–343. Available at: https://doi.org/https://doi.org/10.1007/s10666-012-9346-y.

Khare, V. *et al.* (2022) "Status of renewable energy in Indian subcontinent," *Cleaner Energy Systems*, 3, p. 100020. Available at: https://doi.org/https://doi.org/10.1016/j.cles.2022.100020.

Kibria, M.A., McManus, D.E. and Bhattacharya, S. (2023) "Options for net zero emissions hydrogen from Victorian lignite. Part 2: Ammonia production," *International Journal of Hydrogen Energy*, 48(95), pp. 37166–37182. Available at: https://doi.org/10.1016/j.ijhydene.2023.06.098.

Kissell, R.L. (2020) *Algorithmic Trading Methods: Applications Using Advanced Statistics, Optimization, and Machine Learning Techniques*. 2nd ed. Academic Press. Available at: https://doi.org/10.1016/C2017-0-03456-0.

Klemm, C., Wiese, F. and Vennemann, P. (2023) "Model-based run-time and memory reduction for a mixed-use multi-energy system model with high spatial resolution," *Applied Energy*, 334, p. 120574. Available at: https://doi.org/10.1016/j.apenergy.2022.120574.

Knorr, L. *et al.* (2025) "Electrification and flexibility of process heat in energy system modelling: A review," *Renewable and Sustainable Energy Reviews*, 216, p. 115698. Available at: https://doi.org/10.1016/j.rser.2025.115698.

Kong, K.G.H. *et al.* (2024) "The P-graph application extension in multi-period P2P energy trading," *Renewable and Sustainable Energy Reviews*, 200, p. 114544. Available at: https://doi.org/10.1016/j.rser.2024.114544.

Kotzur, L. *et al.* (2021) "A modeler's guide to handle complexity in energy systems optimization," *Advances in Applied Energy*, 4, p. 100063. Available at: https://doi.org/10.1016/j.adapen.2021.100063.

Koundinya, S. and Seshadri, S. (2022) "Energy, exergy, environmental, and economic (4E) analysis and selection of best refrigerant using TOPSIS method for industrial heat pumps," *Thermal Science and Engineering Progress*, 36, p. 101491. Available at: https://doi.org/10.1016/j.tsep.2022.101491.

Krishnan, A. and Greed, A. (2023) *Energy End Use Database: Key insights and methodology*. Wellington, New Zealand: Energy Efficiency and Conservation Authority. Available at: https://www.eeca.govt.nz/assets/EECA-Resources/Research-papers-guides/EEUD-Key-insights-and-methodology-March-2023.pdf.

Krzemień, J. (2013) "Application of Markal Model Generator in Optimizing Energy Systems," *Journal of Sustainable Mining*, 12, pp. 35–39. Available at: https://doi.org/10.7424/jsm130205.

Kuldna, P., Peterson, K. and Kuhi-Thalfeldt, R. (2015) "Knowledge brokering on emissions modelling in Strategic Environmental Assessment of Estonian energy policy with special reference to the LEAP model," *Environmental Impact Assessment Review*, 54, pp. 55–60.

Kydes, A.S. (1980) "The Brookhaven Energy System Optimization Model: Its Variants and Uses," in W.T. Ziemba and S.L. Schwartz (eds.) *Energy Policy Modeling: United States and Canadian Experiences: Volume II Integrative Energy Policy Models*. Dordrecht: Springer Netherlands, pp. 110–136. Available at: https://doi.org/10.1007/978-94-009-8751-7_7.

Labriet, M. *et al.* (2015) "Assessing Climate Impacts on the Energy Sector with TIAM-WORLD: Focus on Heating and Cooling and Hydropower Potential," in G. Giannakidis et al. (eds.) *Informing Energy and Climate Policies Using Energy Systems Models: Insights from Scenario Analysis Increasing the Evidence Base*. Cham: Springer International Publishing, pp. 389–409. Available at: https://doi.org/10.1007/978-3-319-16540-0_21.

Lahijani, A.M., Protheroe, M.D. and Gschwendtner, M. (2025) “A modelling tool selection for decarbonising industrial process heat systems,” *Renewable and Sustainable Energy Reviews*, 210, p. 115149. Available at: https://doi.org/10.1016/j.rser.2024.115149.

Lambert, J. *et al.* (2023) “Evaluation of sector-coupled energy systems using different foresight horizons,” *Renewable and Sustainable Energy Reviews*, 184, p. 113562. Available at: https://doi.org/10.1016/j.rser.2023.113562.

Lazard (2024) *Lazard’s Levelized Cost of Energy+*. Industry Report. New York, NY: Lazard. Available at: https://www.lazard.com/media/xemfey0k/lazards-lcoeplus-june-2024-_vf.pdf.

Lee, C.H. *et al.* (2020) “A P-graph approach for the synthesis of national-wide bio-hydrogen network from palm oil mill effluent,” *International Journal of Hydrogen Energy*, 45(35), pp. 17200–17219. Available at: https://doi.org/10.1016/j.ijhydene.2020.04.179.

Lenz, V. *et al.* (2020) “Status and Perspectives of Biomass Use for Industrial Process Heat for Industrialized Countries,” *Chemical Engineering & Technology*, 43(8), pp. 1469–1484. Available at: https://doi.org/10.1002/ceat.202000077.

Li, J. *et al.* (2021) “Unit-level cost-benefit analysis for coal power plants retrofitted with biomass co-firing at a national level by combined GIS and life cycle assessment,” *Applied Energy*, 285, p. 116494. Available at: https://doi.org/10.1016/j.apenergy.2021.116494.

Liebreich, M. (2025) *Hydrogen Ladder Version 5.0*. Available at: https://www.linkedin.com/pulse/hydrogen-ladder-version-50-michael-liebreich/ (Accessed: February 25, 2026).

Liew, P.Y. *et al.* (2017) “Total Site Heat Integration planning and design for industrial, urban and renewable systems,” *Renewable and Sustainable Energy Reviews*, 68, pp. 964–985. Available at: https://doi.org/10.1016/j.rser.2016.05.086.

Lin, Y. *et al.* (2025) “Advancing AI-enabled techniques in energy system modeling: a review of data-driven, mechanism-driven, and hybrid modeling approaches,” *Energies*, 18(4), p. 845.

Linnhoff, B. and Flower, J.R. (1978) “Synthesis of heat exchanger networks: I. Systematic generation of energy optimal networks,” *AIChE Journal*, 24(4), pp. 633–642. Available at: https://doi.org/https://doi.org/10.1002/aic.690240411.

LINZ (2026a) *LINZ Data Service*. Available at: https://data.linz.govt.nz/ (Accessed: February 25, 2026).

LINZ (2026b) *New Zealand Transverse Mercator 2000 (NZTM2000)*. Available at: https://www.linz.govt.nz/guidance/geodetic-system/coordinate-systems-used-new-zealand/projections/new-zealand-transverse-mercator-2000-nztm2000 (Accessed: February 25, 2026).

Lipiäinen, S. *et al.* (2021) “Pulp and paper industry in energy transition: Towards energy-efficient and low carbon operation in Finland and Sweden,” *Sustainable Production and Consumption*, 29, pp. 421–431. Available at: https://doi.org/10.1016/j.spc.2021.10.029.

Liu, Z. *et al.* (2020) “An experimental investigation into the effect of flue gas recirculation on ash deposition and Na migration behaviour in circulating fluidized bed during combustion of high sodium Zhundong lignite,” *Fuel Processing Technology*, 199, p. 106300. Available at: https://doi.org/10.1016/j.fuproc.2019.106300.

Löffler, K. *et al.* (2019) “Modeling the low-carbon transition of the European energy system - A quantitative assessment of the stranded assets problem,” *Energy Strategy Reviews*, 26, p. 100422. Available at: https://doi.org/10.1016/j.esr.2019.100422.

Lopion, P. *et al.* (2018) “A review of current challenges and trends in energy systems modeling,” *Renewable and Sustainable Energy Reviews*, 96, pp. 156–166. Available at: https://doi.org/10.1016/j.rser.2018.07.045.

Lopion, P. *et al.* (2019) “Cost Uncertainties in Energy System Optimization Models: A Quadratic Programming Approach for Avoiding Penny Switching Effects,” *Energies*, 12(20). Available at: https://doi.org/10.3390/en12204006.

Lougee-Heimer, R. (2003) “The Common Optimization INterface for Operations Research: Promoting open-source software in the operations research community,” *IBM Journal of Research and Development*, 47(1), pp. 57–66.

Lovegrove, K. *et al.* (2019) *Renewable energy options for industrial process heat*. Turner, ACT, Australia: ITP Thermal Pty Ltd.

Lumbreras, S. and Ciller, P. (2025) “Interpretable optimization: why and how we should explain optimization models,” *Applied Sciences*, 15(10), p. 5732.

Madeddu, S. *et al.* (2020) “The CO2 reduction potential for the European industry via direct electrification of heat supply (power-to-heat),” *Environmental Research Letters*, 15(12), p. 124004. Available at: https://doi.org/10.1088/1748-9326/abbd02.

Maier, R. *et al.* (2025) “Impact of foresight horizons on energy system decarbonization pathways,” *Advances in Applied Energy*, 18, p. 100217. Available at: https://doi.org/10.1016/j.adapen.2025.100217.

Makhorin, A. (2012) “GLPK (GNU Linear Programming Kit).” Available at: https://www.gnu.org/software/glpk/#documentation.

Malehmirchegini, L. and Chapman, A.J. (2025) “Strategies for achieving carbon neutrality within the chemical industry,” *Renewable and Sustainable Energy Reviews*, 217, p. 115762. Available at: https://doi.org/10.1016/j.rser.2025.115762.

Maleki, N. *et al.* (2024) “Future energy insights: Time-series and deep learning models for city load forecasting,” *Applied Energy*, 374, p. 124067.

Malico, I. *et al.* (2019) “Current status and future perspectives for energy production from solid biomass in the European industry,” *Renewable and Sustainable Energy Reviews*, 112, pp. 960–977. Available at: https://doi.org/10.1016/j.rser.2019.06.022.

Manning, C.P. and Fruehan, R.J. (2001) “Emerging technologies for iron and steelmaking,” *JOM*, 53(10), pp. 36–43. Available at: https://doi.org/10.1007/s11837-001-0054-3.

Marina, A. *et al.* (2021) “An estimation of the European industrial heat pump market potential,” *Renewable and Sustainable Energy Reviews*, 139, p. 110545. Available at: https://doi.org/10.1016/j.rser.2020.110545.

Markewitz, P. *et al.* (1996) “IKARUS - A fundamental concept for national GHG-mitigation strategies,” *Energy*, 21(7–8), pp. 693–702. Available at: https://doi.org/10.1016/0196-8904(95)00255-3.

Martí, R., Sevaux, M. and Sörensen, K. (2025) “Fifty years of metaheuristics,” *European Journal of Operational Research*, 321(2), pp. 345–362. Available at: https://doi.org/10.1016/j.ejor.2024.04.004.

Martínez-Gordón, R. *et al.* (2021) “A review of the role of spatial resolution in energy systems modelling: Lessons learned and applicability to the North Sea region,” *Renewable and Sustainable Energy Reviews*, 141, p. 110857. Available at: https://doi.org/10.1016/j.rser.2021.110857.

Mashayekh, S. *et al.* (2017) “A mixed integer linear programming approach for optimal DER portfolio, sizing, and placement in multi-energy microgrids,” *Applied Energy*, 187, pp. 154–168. Available at: https://doi.org/10.1016/j.apenergy.2016.11.020.

Mayer, P. *et al.* (2024) “Flexibility from industrial demand-side management in net-zero sector-coupled national energy systems,” *Frontiers in Energy Research*, 12. Available at: https://doi.org/10.3389/fenrg.2024.1443506.

MBIE (2024a) *Electricity cost and price monitoring*. Available at: https://www.mbie.govt.nz/building-and-energy/energy-and-natural-resources/energy-statistics-and-modelling/energy-statistics/energy-prices/electricity-cost-and-price-monitoring/ (Accessed: February 20, 2026).

MBIE (2024b) *Energy in New Zealand 2024*. Ministry of Business, Innovation and Employment (MBIE). Available at: https://www.mbie.govt.nz/assets/energy-in-nz-2024.pdf.

MBIE (2025a) “Energy in New Zealand.” Available at: https://www.mbie.govt.nz/building-and-energy/energy-and-natural-resources/energy-statistics-and-modelling/energy-publications-and-technical-papers/energy-in-new-zealand.

MBIE (2025b) *Energy in New Zealand 2025*. Ministry of Business, Innovation and Employment (MBIE).

McRae, S. (2025) "Rethinking wholesale market design for New Zealand's clean energy transition," *The Electricity Journal*, p. 107492. Available at: https://doi.org/10.1016/j.tej.2025.107492.

Mercury NZ Limited (2026) "Hydro Generation." Available at: https://www.mercury.co.nz/about-us/renewable-energy/hydro-generation.

Messner, S. and Schrattenholzer, L. (2000) "MESSAGE–MACRO: linking an energy supply model with a macroeconomic module and solving it iteratively," *Energy*, 25(3), pp. 267–282. Available at: https://doi.org/10.1016/S0360-5442(99)00063-8.

Methanex New Zealand (2019) "Submission: Process Heat in New Zealand: Opportunities and Barriers to lowering emissions." Wellington, New Zealand: Ministry of Business, Innovation and Employment (MBIE), Energy Markets.

MetService (2025) *Taupo Weather Station Data*. Available at: https://www.metservice.com/ (Accessed: February 25, 2026).

Milan, C. *et al.* (2015) "Modeling of non-linear CHP efficiency curves in distributed energy systems," *Applied Energy*, 148, pp. 334–347. Available at: https://doi.org/10.1016/j.apenergy.2015.03.053.

Ministry for the Environment (2024) "National Policy Statement for Greenhouse Gas Emissions from Industrial Process Heat 2023." Available at: https://environment.govt.nz/publications/national-policy-statement-for-greenhouse-gas-emissions-from-industrial-process-heat-2023/.

Ministry for the Environment NZ (2024) "New Zealand Land Use Map (LUCAS)." Available at: https://environment.govt.nz/facts-and-science/science-and-data/new-zealand-land-use-map/ (Accessed: February 25, 2026).

Ministry for the Environment NZ (2025) *MfE Data Service*. Available at: https://data.mfe.govt.nz/ (Accessed: February 25, 2026).

Ministry of Transport NZ (2024) *Biomass Transport Analysis*. Te Uru Rākau – New Zealand Forest Service, Ministry for Primary Industries. Available at: https://www.transport.govt.nz/assets/Uploads/DTCC-WP-C5-Road-VOC-June-2023.pdf.

Montanaro, C. *et al.* (2023) "Natural processes leading to large, pre-historic hydrothermal eruptions in geothermal areas: Rotokawa geothermal field, New Zealand," *GSA Bulletin*, 136(1–2), pp. 27–56. Available at: https://doi.org/10.1130/B36509.1.

Morrison, D.R., Sewell, E.C. and Jacobson, S.H. (2014) "An application of the branch, bound, and remember algorithm to a new simple assembly line balancing dataset," *European Journal of Operational Research*, 236(2), pp. 403–409. Available at: https://doi.org/10.1016/j.ejor.2013.11.033.

Mugarza, X. *et al.* (2024) *Industrial Electrification: Strategies and Policies for Europe*. Geneva: World Economic Forum, p. 14. Available at: https://reports.weforum.org/docs/WEF_Industrial_Electrification_2024.pdf.

Muhumuza, R. and Eames, P. (2022) "Decarbonisation of heat: Analysis of the potential of low temperature waste heat in UK industries," *Journal of Cleaner Production*, 372, p. 133759. Available at: https://doi.org/10.1016/j.jclepro.2022.133759.

Münnich, P. *et al.* (2022) *Power-2-Heat: Natural gas savings and climate protection in industry*. 269/05-I-2022/DE. Agora Industrie and FutureCamp, p. 52.

Naegler, T. *et al.* (2016) "Potenziale für erneuerbare Energien in der industriellen Wärmeerzeugung," *BWK- Energie-Fachmagazin*, 68, pp. 20–24.

Nägeli, C. *et al.* (2020) "Towards agent-based building stock modeling: Bottom-up modeling of long-term stock dynamics affecting the energy and climate impact of building stocks," *Energy and Buildings*, 211, p. 109763. Available at: https://doi.org/10.1016/j.enbuild.2020.109763.

Naqvi, M., Yan, J. and Dahlquist, E. (2010) "Black liquor gasification integrated in pulp and paper mills: A critical review," *Bioresource Technology*, 101(21), pp. 8001–8015. Available at: https://doi.org/10.1016/j.biortech.2010.05.013.

Narayanan, M. *et al.* (2020) "Development of a Coupled TRNSYS-MATLAB Simulation Framework for Model Predictive Control of Integrated Electrical and Thermal Residential Renewable Energy System," *Energies*, 13(21). Available at: https://doi.org/10.3390/en13215761.

Neniškis, E. and Galinis, A. (2018) "Representation of wind power generation in economic models for long-term energy planning," *Energetika*, 64. Available at: https://doi.org/10.6001/energetika.v64i1.3726.

Neurath, O. (1973) "From Vienna Method to ISOTYPE," in M. Neurath and R.S. Cohen (eds.) *Empiricism and Sociology*. Dordrecht: Reidel, p. 220.

New Zealand Meat Board (2026) "New Zealand Meat Board." Level 4, Wellington Chambers, 154 Featherston Street, Wellington, New Zealand. Available at: https://www.nzmeatboard.org/.

NIWA (2025) *NIWA*. Available at: https://niwa.co.nz/ (Accessed: February 25, 2026).

Noorollahi, Y. *et al.* (2021) "Modeling of Transition from Natural Gas to Hybrid Renewable Energy Heating System," *International Journal of Sustainable Energy Planning and Management*, 32, pp. 61–78. Available at: https://doi.org/10.5278/ijsepm.6576.

Obeng, M. *et al.* (2019) "Technical and economic feasibility of a 50MW Grid-connected Solar PV at UENR Nsoatre Campus," *Journal of Cleaner Production*, 247, p. 119159. Available at: https://doi.org/10.1016/j.jclepro.2019.119159.

Oberschelp, C., Pfister, S. and Hellweg, S. (2023) “Global site-specific health impacts of fossil energy, steel mills, oil refineries and cement plants,” *Scientific Reports*, 13(1), p. 13708. Available at: https://doi.org/10.1038/s41598-023-38075-z.

Olsson, O. and Schipfer, F. (2021) *Decarbonizing industrial process heat: the role of biomass*. IEA Bioenergy.

Open Country Dairy (2026) “Sustainability.” Available at: https://www.opencountry.co.nz/en-nz/about-us/sustainability/.

Open Energy Platform (2025) *Open Energy Platform*. Available at: https://openenergyplatform.org/ (Accessed: February 25, 2026).

Openmod Initiative (2025) *Openmod Initiative*. Available at: https://openmod-initiative.org/ (Accessed: February 25, 2026).

Pag, F. *et al.* (2024) “Oversizing solar heating plants in industry: A cost-effective solution to increase solar fractions,” *Solar Energy Advances*, 4, p. 100071. Available at: https://doi.org/10.1016/j.seja.2024.100071.

Papapetrou, M. *et al.* (2018) “Industrial waste heat: Estimation of the technically available resource in the EU per industrial sector, temperature level and country,” *Applied Thermal Engineering*, 138, pp. 207–216. Available at: https://doi.org/10.1016/j.applthermaleng.2018.04.043.

Payet-Burin, R. *et al.* (2021) “The Impact of Assuming Perfect Foresight When Planning Infrastructure in the Water–Energy–Food Nexus,” *Frontiers in Water*, 3, p. 778003. Available at: https://doi.org/https://doi.org/10.3389/frwa.2021.778003.

Pedro, H.T.C. *et al.* (2018) “Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts,” *Renewable Energy*, 123, pp. 191–203. Available at: https://doi.org/10.1016/j.renene.2018.02.006.

Pfenninger, S. (2017) “Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability,” *Applied Energy*, 197, pp. 1–13. Available at: https://doi.org/10.1016/j.apenergy.2017.03.051.

Pfenninger, S. (2024) “Open code and data are not enough: Understandability as design goal for energy system models,” *Progress in Energy*, 6(3), p. 033002. Available at: https://doi.org/10.1088/2516-1083/ad371e.

Pfenninger, S., Hawkes, A. and Keirstead, J. (2014) “Energy systems modeling for twenty-first century energy challenges,” *Renewable and Sustainable Energy Reviews*, 33, pp. 74–86. Available at: https://doi.org/10.1016/j.rser.2014.02.003.

Pfenninger, S. and Staffell, I. (2016) “Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data,” *Energy*, 114, pp. 1251–1265. Available at: https://doi.org/10.1016/j.energy.2016.08.060.

Philibert, C. (2017) *Renewable Energy for Industry: From Green Energy to Green Materials and Fuels*. Paris: International Energy Agency. Available at: https://www.iea.org/reports/renewable-energy-for-industry.

Phillips, K. *et al.* (2023) “Spatial representation of renewable technologies in generation expansion planning models,” *Applied Energy*, 342, p. 121092. Available at: https://doi.org/10.1016/j.apenergy.2023.121092.

Philpott, A. *et al.* (2019) “The New Zealand Electricity Market: Challenges of a Renewable Energy System,” *IEEE Power and Energy Magazine*, 17, pp. 43–52. Available at: https://doi.org/10.1109/MPE.2018.2871705.

Poletti, S. (2021) “Market Power in the New Zealand electricity wholesale market 2010–2016,” *Energy Economics*, 94, p. 105078. Available at: https://doi.org/10.1016/j.eneco.2020.105078.

Poluzzi, A., Guandalini, G. and Romano, M. (2022) “Flexible methanol and hydrogen production from biomass gasification with negative emissions,” *Sustainable Energy & Fuels*, 6. Available at: https://doi.org/10.1039/D2SE00661H.

Pöstges, A. and Weber, C. (2023) “Identifying key elements for adequate simplifications of investment choices – The case of wind energy expansion,” *Energy Economics*, 120, p. 106534. Available at: https://doi.org/10.1016/j.eneco.2023.106534.

Pratschner, S. *et al.* (2023) “Techno-economic assessment of a power-to-green methanol plant,” *Journal of CO2 Utilization*, 75, p. 102563. Available at: https://doi.org/10.1016/j.jcou.2023.102563.

Prestigiacomo, C. *et al.* (2024) “Solar heat for the decarbonization of chemical industry: dehydrogenation of ethylbenzene to styrene driven by a concentrating solar power plant with molten salts as heat transfer fluids,” *Energy Conversion and Management: X*, 22, p. 100546. Available at: https://doi.org/10.1016/j.ecmx.2024.100546.

Rahmaniani, R. *et al.* (2017) “The Benders decomposition algorithm: A literature review,” *European Journal of Operational Research*, 259(3), pp. 801–817. Available at: https://doi.org/10.1016/j.ejor.2016.12.005.

Rajaee, F. *et al.* (2024) “Techno-economic evaluation of biomass-to-methanol production via circulating fluidized bed gasifier and solid oxide electrolysis cells: A comparative study,” *Energy Conversion and Management*, 301, p. 118009. Available at: https://doi.org/10.1016/j.enconman.2023.118009.

Raman, R. and Grossmann, I.E. (1994) “Modelling and computational techniques for logic based integer programming,” *Computers & Chemical Engineering*, 18(7), pp. 563–578. Available at: https://doi.org/10.1016/0098-1354(93)E0010-7.

Ramsebner, J. *et al.* (2021) “The sector coupling concept: A critical review,” *WIREs Energy and Environment*, 10(4), p. e396. Available at: https://doi.org/https://doi.org/10.1002/wene.396.

Reid, G. *et al.* (2025) “Modelling Future Pathways for Industrial Process Heat Decarbonisation in New Zealand: The Role of Green Hydrogen,” *Sustainability*, 17(23). Available at: https://doi.org/10.3390/su172310812.

Ridha, E., Nolting, L. and Praktiknjo, A. (2020) “Complexity profiles: A large-scale review of energy system models in terms of complexity,” *Energy Strategy Reviews*, 30, p. 100515. Available at: https://doi.org/10.1016/j.esr.2020.100515.

Ringkjøb, H.-K., Haugan, P. and Solbrekke, I. (2018) “A review of modelling tools for energy and electricity systems with large shares of variable renewables,” *Renewable and Sustainable Energy Reviews*, 96, pp. 440–459. Available at: https://doi.org/10.1016/j.rser.2018.08.002.

Rissman, J. *et al.* (2020) “Technologies and policies to decarbonize global industry: Review and assessment of mitigation drivers through 2070,” *Applied Energy*, 266, p. 114848. Available at: https://doi.org/10.1016/j.apenergy.2020.114848.

Rosa, L. and Castro, R. (2020) “Forecasting and assessment of the 2030 australian electricity mix paths towards energy transition,” *Energy*, 205, p. 118020. Available at: https://doi.org/10.1016/j.energy.2020.118020.

Rossetti, I. *et al.* (2023) “Conceptual design of a process for hydrogen production from waste biomass and its storage in form of liquid ammonia,” *International Journal of Hydrogen Energy*, 48(96), pp. 37443–37460. Available at: https://doi.org/10.1016/j.ijhydene.2023.01.261.

Roth, M., Franke, G. and Rinderknecht, S. (2023) “Optimal component sizing and operational optimisation of a mobile energy system for decentralised electricity supply,” *Smart Energy*, 11, p. 100108. Available at: https://doi.org/10.1016/j.segy.2023.100108.

Rubinsin, N.J. *et al.* (2024) “An overview of the enhanced biomass gasification for hydrogen production,” *International Journal of Hydrogen Energy*, 49, pp. 1139–1164. Available at: https://doi.org/10.1016/j.ijhydene.2023.09.043.

Salehmin, M.N.I. *et al.* (2024) “Navigating challenges and opportunities of machine learning in hydrogen catalysis and production processes: Beyond algorithm development,” *Journal of Energy Chemistry*, 99, pp. 223–252.

Schoeneberger, C. *et al.* (2022) “Electrification potential of U.S. industrial boilers and assessment of the GHG emissions impact,” *Advances in Applied Energy*, 5, p. 100089. Available at: https://doi.org/10.1016/j.adapen.2022.100089.

Sensfuß, F., Ragwitz, M. and Genoese, M. (2008) “The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany,” *Energy Policy*, 36(8), pp. 3086–3094. Available at: https://doi.org/10.1016/j.enpol.2008.03.035.

Shah, S. (2025) “Monte Carlo Simulation in Renewable Energy Planning: A Comprehensive Review and Novel Framework for Uncertainty Quantification,” *The American Journal of Engineering and Technology*, 7(06), pp. 24–45. Available at: https://doi.org/10.37547/tajet/Volume07Issue06-04.

Shahabuddin, M., Brooks, G. and Rhamdhani, M.A. (2023) “Decarbonisation and hydrogen integration of steel industries: Recent development, challenges and technoeconomic analysis,” *Journal of Cleaner Production*, 395, p. 136391. Available at: https://doi.org/10.1016/j.jclepro.2023.136391.

Siskos, P. *et al.* (2018) “Implications of delaying transport decarbonisation in the EU: A systems analysis using the PRIMES model,” *Energy Policy*, 121. Available at: https://doi.org/10.1016/j.enpol.2018.06.016.

Skolfield, K. *et al.* (2025) “IDAES GTEP 0.1 Release.” Available at: https://doi.org/10.18141/2568455.

Slimani, J. *et al.* (2024) *Towards a Sustainable Energy Future: Modeling Morocco’s Transition to Renewable Power with Enhanced Osemosys Model*. Available at: https://doi.org/10.2139/ssrn.4826749.

Stats NZ (2024) *Stats NZ*. Available at: https://datafinder.stats.govt.nz/.

Stein, D. *et al.* (2017) *Development of adaptive time patterns for multi-dimensional power system simulations*, p. 5. Available at: https://doi.org/10.1109/EEM.2017.7981868.

Sun, S. (2013) “A review of deterministic approximate inference techniques for Bayesian machine learning,” *Neural Computing and Applications*, 23(7), pp. 2039–2050. Available at: https://doi.org/10.1007/s00521-013-1445-4.

Sweeney, J. (1981) “World oil. Preliminary results from the energy modelling forum-study,” *Kvartalsskrift, Bergen Bank;(Norway)*, 3.

Tan, R.R. and Aviso, K.B. (2016) “An extended P-graph approach to process network synthesis for multi-period operations,” *Computers & Chemical Engineering*, 85, pp. 40–42. Available at: https://doi.org/10.1016/j.compchemeng.2015.10.007.

Tatit, P., Adhinugraha, K. and Taniar, D. (2024) “Navigating the Maps: Euclidean vs. Road Network Distances in Spatial Queries,” *Algorithms*, 17(1). Available at: https://doi.org/10.3390/a17010029.

Tattini, J. *et al.* (2018) “A Long-Term Strategy to Decarbonise the Danish Inland Passenger Transport Sector,” in G. Giannakidis et al. (eds.) *Limiting Global Warming to Well Below 2 °C: Energy System Modelling and Policy Development*. Cham: Springer International Publishing, pp. 137–153. Available at: https://doi.org/10.1007/978-3-319-74424-7_9.

Tawarmalani, M. and Sahinidis, N.V. (2002) *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications*. New York, NY: Springer (Nonconvex Optimization and Its Applications). Available at: https://doi.org/10.1007/978-1-4757-3532-1.

Taylor, M.T., Atkins, M.J. and Walmsley, T.G. (2025) "Hybrid industrial utility systems with biomass and electrode boilers: a techno-economic investigation," *Renewable and Sustainable Energy Transition*, 7, p. 100116. Available at: https://doi.org/10.1016/j.rset.2025.100116.

Teng, S.Y. *et al.* (2022) "Framework to embed machine learning algorithms in P-graph: Communication from the chemical process perspectives," *Chemical Engineering Research and Design*, 188, pp. 265–270. Available at: https://doi.org/10.1016/j.cherd.2022.09.043.

von Thadden del Valle, C. *et al.* (2024) "Technologies for defossilisation of process heat: An analysis and overview of industrial processes and temperature requirements," *Proceedings of the 37th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems*. Rhodes, Greece, p. 1.

Thiel, G.P. and Stark, A.K. (2021) "To decarbonize industry, we must decarbonize heat," *Joule*, 5(3), pp. 531–550. Available at: https://doi.org/10.1016/j.joule.2020.12.007.

Thomsen, J. *et al.* (2021) "Effect of the Foresight Horizon on Computation Time and Results Using a Regional Energy Systems Optimization Model," *Energies*, 14(2). Available at: https://doi.org/10.3390/en14020495.

Tollefson, J. (2017) "The wooden skyscrapers that could help to cool the planet," *Nature*, 545(7654), pp. 280–282. Available at: https://doi.org/10.1038/545280a.

Touafio, J. *et al.* (2020) "Analysis of a wind turbine project in the city of Bouar (Central African Republic)," *Scientific African*, 8, p. e00354. Available at: https://doi.org/10.1016/j.sciaf.2020.e00354.

Transpower New Zealand (2022a) *Maps and GIS data*. Available at: https://www.transpower.co.nz/our-work/industry/our-grid/maps-and-gis-data (Accessed: December 28, 2025).

Transpower New Zealand (2022b) *Transpower*. Available at: https://www.transpower.co.nz/ (Accessed: February 20, 2024).

Transpower New Zealand (2023a) *Transmission Planning Report 2023*. Transpower New Zealand. Available at: https://static.transpower.co.nz/public/uncontrolled_docs/Transmission%20Planning%20Report%202023.pdf.

Transpower New Zealand (2023b) *Winter Peak Analysis: 2024 and 2025*. Transpower New Zealand Limited. Available at: https://static.transpower.co.nz/public/bulk-upload/documents/Winter%20Peak%20Analysis%202024-25.pdf.

Transpower New Zealand (2025a) “TPR Envision and Envision Opportunities.” Available at: https://experience.arcgis.com/experience/007dcdef2909420ba7420377a22799d0/page/TPR-ENVISION.

Transpower New Zealand (2025b) “Transpower Open Data Portal.” Available at: https://data-transpower.opendata.arcgis.com/.

Vielma, J.P. (2015) “Mixed Integer Linear Programming Formulation Techniques,” *SIAM Review*, 57(1), pp. 3–57. Available at: https://doi.org/10.1137/130915303.

Villemin, T. *et al.* (2024) “Monte Carlo prediction of the energy performance of a photovoltaic panel using detailed meteorological input data,” *International Journal of Thermal Sciences*, 195, p. 108672. Available at: https://doi.org/10.1016/j.ijthermalsci.2023.108672.

Visser, H., Thopil, G.A. and Brent, A. (2019) “Life cycle cost profitability of biomass power plants in South Africa within the international context,” *Renewable Energy*, 139, pp. 9–21. Available at: https://doi.org/10.1016/j.renene.2019.02.080.

Wadström, C., Johansson, M. and Wallén, M. (2021) “A framework for studying outcomes in industrial symbiosis,” *Renewable and Sustainable Energy Reviews*, 151, p. 111526. Available at: https://doi.org/10.1016/j.rser.2021.111526.

Walden, J.V.M., Wellig, B. and Stathopoulos, P. (2023) “Heat pump integration in non-continuous industrial processes by Dynamic Pinch Analysis Targeting,” *Applied Energy*, 352, p. 121933. Available at: https://doi.org/10.1016/j.apenergy.2023.121933.

Wang, Q. *et al.* (2023) “Enhancing Energy Transition through Sector Coupling: A Review of Technologies and Models,” *Energies*, 16(13). Available at: https://doi.org/10.3390/en16135226.

Xu, Y. *et al.* (2023) “Optimal renewable energy export strategies of islands: Hydrogen or electricity?,” *Energy*, 269, p. 126750. Available at: https://doi.org/10.1016/j.energy.2023.126750.

Yue, X. *et al.* (2018) “From 2 °C to 1.5 °C: How Ambitious Can Ireland Be?,” in G. Giannakidis et al. (eds.) *Limiting Global Warming to Well Below 2 °C: Energy System Modelling and Policy Development*. Cham: Springer International Publishing, pp. 191–205. Available at: https://doi.org/10.1007/978-3-319-74424-7_12.

Zhang, S. and Xie, L. (2023) “Grafting constructive algorithm in feedforward neural network learning,” *Applied Intelligence*, 53(10), pp. 11553–11570. Available at: https://doi.org/10.1007/s10489-022-04082-2.

Zhao, Q. *et al.* (2019) “Energy Flows and Carbon Footprint in the Forestry-Pulp and Paper Industry,” *Forests*, 10(9). Available at: https://doi.org/10.3390/f10090725.

Zou, J., Ahmed, S. and Sun, X. (2018) “Stochastic dual dynamic integer programming,” *Mathematical Programming*, 175. Available at: https://doi.org/10.1007/s10107-018-1249-5.

Zühlsdorf, B. *et al.* (2023) *IEA HPT Annex 58: High-Temperature Heat Pumps. Task 1 report: Technologies*. IEA Heat Pumping Technologies. Available at: https://heatpumpingtechnologies.org/annex58/wp-content/uploads/sites/70/2023/09/annex-58-task-1-technologies-task-report.pdf.

Appendix A

Regular Expression for the P-graph Output File

As for the output text file, the overview where data is extracted mainly in the “Materials”, “Operating Units and Flow” and “Total annual cost” section as shown in Figure A 1. The Materials and Operating Units list at the beginning of the text shows the nodes fed into the optimisation run and the maximal structure shows that the nodes which are in the combinatorially feasible process structure for before a solution is obtained. Overall, it is obtainable to extract values from the Materials section whereas the Operating Units and Flow section would require more steps. This step combines multi-line operating unit descriptions into a single logical row, for example merging the input and output lines of an operating unit into one continuous record.

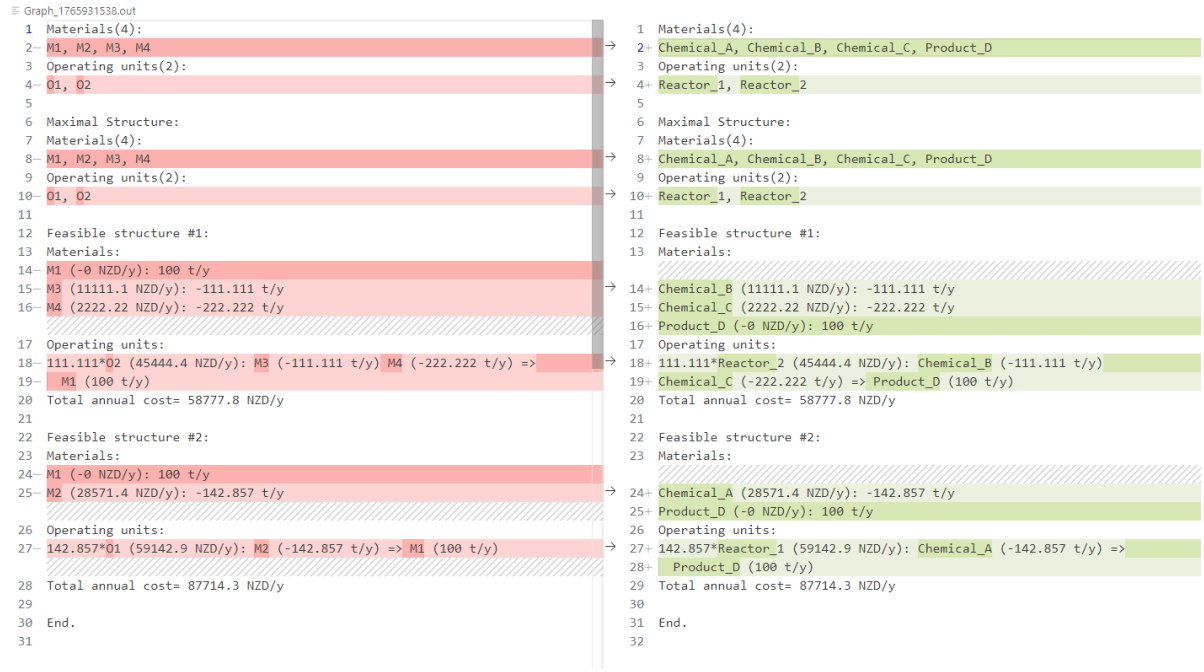


Figure A 1: Different formatting for the output files for P-graph Python (left) generated and P-graph Studio (right).

While making the regex flexible and usable for both formats for the Material Section, Rule 2 and Rule 9 are applied to omit possible blank spaces which would affect the parsing, while also using some keeping some literal text such as brackets “()” capture the group for cost value and cost unit for each row. Literal text of symbol semicolon “:” is also added to segregate cost information with numeric information for each material node. The table shows the key regex used and their application purpose in the Table A 1. In order to understand how each regex rule is applied, the example is shown in Figure A 2. A brief description to extract both materials nodes are shown as follows.

Table A 1: General Expression Rule for the Material Flow Balance.

Description	Group
Material ID or Material name	(.+?)
Cost Value in Numeric	(-?[\d\.eE\+\-]+)
Cost Unit	([^\)]+)
Values accumulated in the material node	(-?[\d\.eE\+\-]+)
Flow unit	(.+)

The regex for parsing each line in the Materials section is shown as:

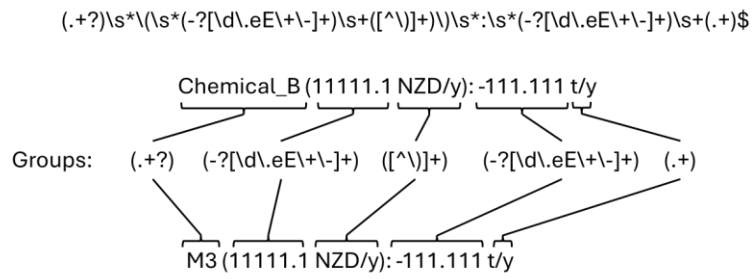


Figure A 2: Regex expression for the material section.

Operating Units and Flow Section

To parse the Operating Units and Flow section and retrieve the capacity multiplier, operating unit name or ID, and annualised cost, the output text is read line by line while stripping unnecessary whitespace. This reduces sensitivity to inconsistent spacing and line breaks.

Step 1 is to consolidate multi-line operating unit descriptions into a single row. Each operating unit node entry can be identified by the presence of an asterisk (*) which denotes the capacity multiplier value and marks the start of a new operating unit node entry. Any subsequent line that does not contain an asterisk is treated as a continuation line of the previous line and appended to the previous record. This creates one complete string per operating unit entry, even when the original output splits information across multiple lines.

In Step 2 is to extract the operating unit attributes from the string. Once all continuation lines have been merged, the completed entries are stored in a list and parsed one by one using regular expressions. A dedicated operating unit regex is applied to extract the capacity multiplier, operating unit identifier, cost value, and cost unit. The regular expressions are displayed in Table A 2.

Table A 2: Regular expression capture groups used to parse operating unit entries from the P-graph output file.

No	Description	Group
1	Capacity Multiplier	([\d\.eE+\-]+)
2	Operating Unit ID or Name	([A-Za-z][\w]*)
3	Cost Value	([\d\.eE+\-]+)
4	Cost Unit	([^\)]+)

For the next step, the flow information from the same consolidated entry is extracted. The flow description is separated into inlet and outlet terms using the "=>" delimiter, where the left side represents flows entering the operating unit and the right side represents flows leaving the operating unit. Each side is then processed using a regular expression "findall" function with greedy matching, because multiple material terms can appear in the same direction within a single line. Table A 3 shows the regular expression used to parse flow terms.

Table A 3: Regular expression capture groups used to parse material flow terms from the P-graph output file.

No	Description	Group
1	Material ID or Name	([A-Za-z]\w*)
2	Flow Value	([+\-]?\d+(?:\.\d+)?(?:[eE][-+]?\d+)?)
3	Flow Unit	([^)]+)

Step 4 is organised to extracted flow terms into directed tuples. Each regular expression rule match captures the values from material identifier, flow value, and flow unit. Inlet flows are being stored as variables named (Material, Operating Unit, Flow, Unit), whereas the outlet flows are stored as (Operating Unit, Material, Flow, Unit). These tuples provide a consistent representation of arc direction and coefficient values, so as to align with how flows are interpreted in the P-graph structure.

Lastly, the extracted operating unit attributes and flow tuples are stored in a dictionary using consistent keys such as "Source", "Destination", "Flow", and "Unit". The P-graph Studio output file is used as the example because its formatting is typically more irregular than the identifier formatted output generated by the P-graph Python workflow, making it a good example for robust testing for parsing the Operating Units and Flow Section. Figure A 3 shows the summary workflow for the Operating Units and Flow Section.

<table>
<tr><th colspan="3">Example from P-graph Studio:</th></tr>
<tr><td colspan="3">Looping through the Lines

Line 26 ……
Line 27: 142.857*Reactor_1 (59142.9 NZD/y): Chemical_A (-142.857 t/y) =>
Line 28: Product_D (100 t/y)
Line 29 ……</td></tr>
<tr><td colspan="3">Step 1: Use an “if” statement to identify full flow and capacity relations by checking whether an asterisk (*) exists in the line.</td></tr>
<tr><td colspan="2">[Line 27]
True
↓
Operating Unit + Flow Entry
↓</td><td>[Line 28]
False
↓
Append to Previous Line
↙</td></tr>
<tr><td colspan="3">Output:

142.857*Reactor_1 (59142.9 NZD/y): Chemical_A (-142.857 t/y) => Product_D (100 t/y)</td></tr>
<tr><td colspan="3">Apply Regular Expression Rule to Split Operating Unit Entries with Flow Entries</td></tr>
<tr><td>Operating Unit
Step 2: Apply RE</td><td colspan="2">Flows
Step 3: Apply RE</td></tr>
<tr><td rowspan="2">Capacity Multiplier: 142.857
Operating Unit ID: Reactor_1
Cost Value: 59142.9
Cost Unit: NZD/y</td><td>Flow 1
Material ID:Chemical_A
Flow Value: 142.857
Flow Unit: t/y</td><td>Flow 2
Material ID:Product_D
Flow Value: 100
Flow Unit: t/y</td></tr>
<tr><td colspan="2">Organise into Flow Tuple follow the format:
(Source, Destination, Flow, Unit)

(Chemical_A, Reactor_1, 142.857, t/y)
(Reactor_1, Product_D, 100, t/y)</td></tr>
</table>

Figure A 3: Summary workflow for the Operating Units and Flow Section.

Appendix B

Generation Example

For the electricity generation layer, identifiers are filtered based on a structured naming convention. Operating unit nodes associated with generation capacity use identifiers with a length of six characters and a prefix of O14. In this convention, O denotes an operating unit node, the digit 1 represents the generation layer, and the digit 4 corresponds to the capacity stream of the generation unit. The remaining digits identify the generation site or company.

To analyse how electricity is allocated between grid export and methanol production, additional filtering rules are applied. Biomass fuel consumption for electricity generation is identified using the prefix O12 with a character length of eight. These identifiers include a time dependent element, allowing biomass use to be tracked across time steps. This enables assessment of competition between biomass consumption for electricity generation and direct industrial heating.

Each generation site may serve both grid export and on-site methanol production, and therefore two masking rules are applied. Electricity supplied to the grid is identified using the prefix O18, followed by the site index and a time dependent suffix. Electricity supplied to methanol production is identified using the prefix O1, reflecting its close association with the generation plant. These identifiers include both a time element and a dedicated flow capacity code, allowing electricity flows to methanol production to be distinguished from grid exports.

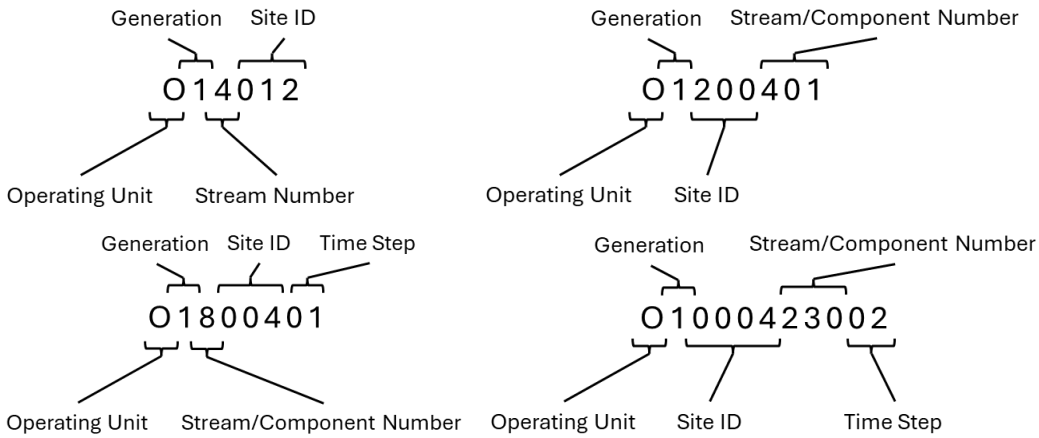


Figure B 1: Illustrates the format on the breakdown identifier for the generation layer.

Methanol Example

A tabular representation based on the methanol plants point of view is also generated in the model. The template of the layer presents two sheets with one presents the capacity built in the optimised structure whereas another presents the green methanol production with time. The sizing of the methanol plant is based on identifier patterns. The capacity identifier has 9 characters, a prefix of O1 and an identifier sixth till eighth location of 280 for the methanol plant. The capacity shows what is the cost optimal of the energy system to fulfil green methanol production

requirements or maximising profits for producing a higher value selling product. With the addition of time element, the identifier has a length of eleven characters, and the sixth till eighth location is changed to 260, this would show what is the production rate of the methanol production for each given time.

Additional templates could be applied, while analysing the perspective of the green methanol plant. As an example, identifying which biomass and type is transported to the methanol plant, as well as the flow and transport cost associated to the respective site.

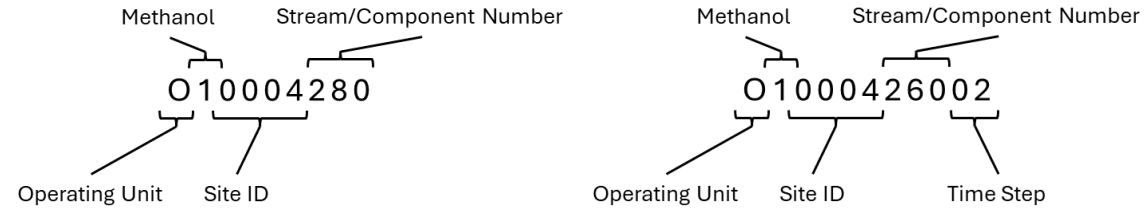


Figure B 2: Illustrates the format on the breakdown identifier for the methanol layer.

Biomass Resource

Key biomass resource tabulation format includes flow of biomass to factories, and biomass resource use limit. These two are the prime indicators on how a spatial analysis is advantageous over an aggregated energy model. From the flow of biomass to factories, specifications such as types of biomasses, travel distance and energy flow of biomass can be solved by the optimisation model. This is particularly important to identify the effect of limited resource type and distance constraint can change the optimal structure of the model. The identifier consists of the append string of both biomass resource, biomass resource type with the respective factory. In the model illustration, the identifier effectively includes the interaction between both layers, where the identifier format helps to also visualise the connection flow as well as storing key attributes such as biomass type which exhibits different heat value, moisture level hence different transport cost per unit energy.

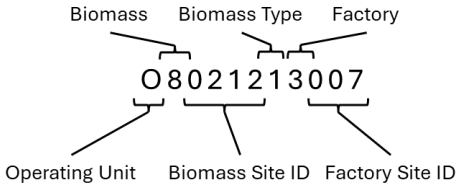


Figure B 3: Illustrates the format on the breakdown identifier for the biomass to factory layer.

From the spatial biomass perspective, when biomass is made as the first header column, the identifier would show how much biomass resource is consumed in the model solution for each resource type at the location. The pattern would show which region is saturated with biomass resources where another region is less reliant on biomass. The masking for the layer would be a prefix of M8 and length of 6. The prefix of M indicates that the value representing the resources lie within the material nodes sheet. Also, since the entry for each identifier is in row, the biomass type is extracted and pivoted into columns for each of analysing the results.

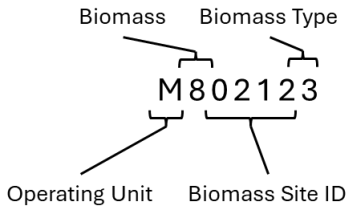


Figure B 4: Illustrates the format on the breakdown identifier for the biomass layer.

Transmission Grid

Looking into the transmission grid, the ID masking layer is length of the character of twelve, starts with O6 and ends with 3 as a unique string component number which represents capacity in the model. The identifier contains the unique matching of the specific index in the input table, where the index represents the rows shows the joining node in the transmission sheet input, this is because the index in every entry in the row has an attribute which traces the POC name where the first append is the source and the second part would be the destination. Line number represents the number of lines of the current identical existing capacity are required to be upgraded to ensure the feasible transmission structure.

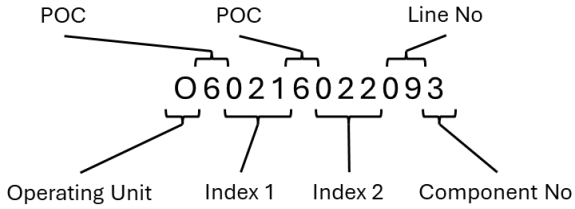


Figure B 5: Illustrates the format on the breakdown identifier for the transmission layer.

Attributes such as POC names, grid name, distance, region, grid capacity and transmission voltage can be merged with the solution structure while dropping the index numbers which are arbitrary, show the additional quantity and capacity of electricity are required for electricity delivery.

Transformer

The transformer layer uses two distinct filtering rules to differentiate between supply and demand transformers. In this convention, transformer related identifiers have a length of eight characters and include a suffix of 53, which denotes transformer capacity. Characters in positions three to five of the identifier specify the transformer type and the associated POC. The asset number has to be present as it considered unique as different transformer types would have different operational capacities.

To distinguish transformer function efficiently, a numerical threshold is applied to the value in positions three to five. Values greater than 300 indicate demand-side transformers, while values less than 300 represent supply side transformers. For demand transformers, the site identifier is obtained by subtracting 300 from this value. This offset serves as a placeholder to identify transformer type within the same identifier format, avoiding the need to introduce additional characters into the identifier string. Accordingly, the example shown in Figure X represents a

demand-side transformer associated with site number 21 and corresponding to transformer number 6.

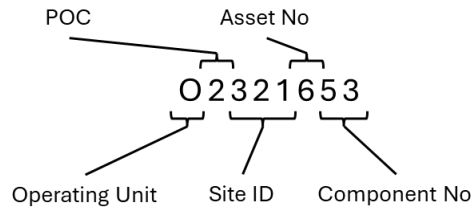


Figure B 6: Illustrates the format on the breakdown identifier for the transformer layer.

Distribution Line

At the distribution level, upgrades to distribution lines are assessed for factories with high electrification uptake. In such cases, additional distribution lines may be installed between factories and their connected POCs to increase electricity throughput within the energy system. These upgrades are represented using prefixes associated with the factory layer data. They should not be confused with electricity consumption within the factory itself, as they represent network infrastructure enhancements required to support fuel-switching for process heat.

The identifier segment located in positions six to eight represents the stream number within this layer and is deliberately assigned the value 103. This design ensures consistency with the electricity inflow stream to each factory, including its time dependent representation, allowing electricity flows associated with distribution upgrades to be tracked coherently across time periods.

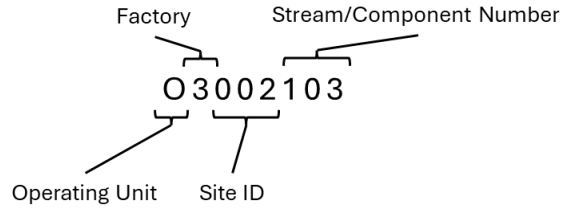


Figure B 7: Illustrates the format on the breakdown identifier for the distribution layer.

Sub Transmission

In the post processing stage for the sub transmission layer, electricity grid upgrades are primarily assessed for newly proposed generation sites, as existing generation facilities typically already have installed grid infrastructure. However, since the capacity of electricity generation plants is modelled as a continuous variable, subject to an upper bound defined by consented capacity limits, the capacities of the connecting lines must be scaled proportionally to the generation output. As a result, sub transmission grid upgrades are closely associated with the generation layer and represent a critical consideration for power companies when planning and financing new connection infrastructure.

To support this analysis, unique identifiers are constructed using prefixes associated with the generation layer, combined with generation site identifiers to distinguish specific line upgrades.

These identifiers also encode information on the line type selection, allowing the model to evaluate which available line configurations are cost effective while respecting the integer sizing constraints of transmission assets available in the market.

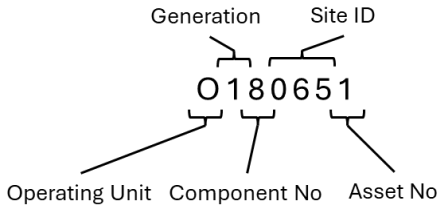


Figure B 8: Illustrates the format on the breakdown identifier for the sub transmission layer.

Interlayer Representation

Interlayer representation enables data to be extracted from multiple system layers and combined into a unified tabular format. One of the primary applications of this approach is the estimation of total electricity demand at each POC. This is achieved by combining projected baseline electricity demand with the aggregated electricity demand arising from newly electrified industrial processes at the same POC. The resulting comparison allows assessment of the increase in electricity demand within the distribution layer, supporting sensitivity checks and highlighting locations that are already heavily reliant on electricity.

This process begins by extracting non-industrial process electricity demand using an identifier-based filtering rule with a prefix of O4 and a total identifier length of seven characters. The site identifier embedded in this code is defined in the total identifier. The final two characters of the identifier represent the time step and are aggregated when annual demand at the POC level is required.

In parallel, electricity flows associated with distribution to factories are extracted using the corresponding distribution layer identifiers, which also include time dependent elements. These flows represent electricity delivered to factories prior to conversion by heating technologies and therefore capture the gross electrical input before accounting for thermal efficiency. Once both baseline and industrial electricity demands are aggregated, the total electricity demand at each POC can be evaluated and analysed.

As electricity flows across multiple layers, it is more effective to identify the unique masking pattern of the predecessor node from which flows branch into successor nodes belonging to different layers. This approach simplifies the aggregation process, as it allows values to be merged based on a common upstream reference rather than relying on identifiers that do not share consistent character patterns.

Once the predecessor node is identified, the corresponding destination nodes that carry the relevant values can be retrieved. These values are then summed and represented at the POC layer, enabling consistent aggregation of electricity flows across multiple layers. The site identifier embedded in the node string is subsequently adjusted by subtracting 300 to recover the actual site number.

Figure B 9 illustrates the relationship between the predecessor identifier format and the successor identifier formats, which do not share common character patterns and therefore cannot be matched directly without this intermediate mapping step.

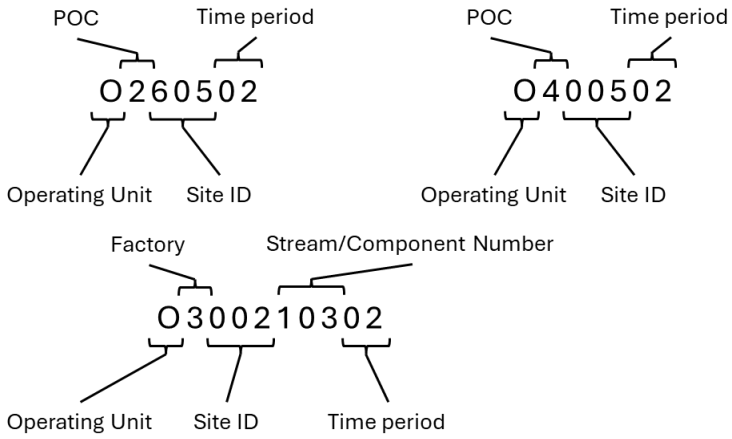


Figure B 9: Illustrates the format on the breakdown identifier for the interlayer layer between the grid networks.

Appendix C

Factory Model

Biomass Resource Layer

Several key equations include biomass resource constraints in the model which is represented as such.

$$0 \leq B_{k,m} \leq \overline{B}_{k,m} y_{k,m}^{bio} \tag{C 1}$$

As biomass is transported between different layer, the balance equation can be observed in Equation (C 2). Connection sets are determined from preprocessing which establish the connections between nodes, showing the geographical feasibility between sites.

$$\sum_{(k,j)\in A^{bio,MeOH}} B_{k,j,m}^{bio,MeOH} + \sum_{e\in P^{th}} \sum_{(k,j)\in A^{bio,gen}} B_{k,e,j,m}^{bio,gen} + \sum_{(k,l)\in A^{bio,fac}} B_{k,l,m}^{bio,fac} = B_{k,m} \quad \forall k,m \tag{C 2}$$

In order to reduce the complexity and total number of variables of the model, the time dimension T, is introduced in Equation (C 3) after the connections have been established. Hence, π_t is introduced with $\sum_{t\in T} \pi_t = 1$ in Equation (C 4) and Equation (C 5).

$$F_{l,m,t}^{bio,fac} = \pi_t \sum_{k\in K} B_{k,l,m}^{bio,fac} \tag{C 3}$$

$$F_{j,t}^{bio,gen} = \pi_t \sum_{e\in P^{th}} \sum_{m\in M} \sum_{k\in K} B_{k,e,j,m}^{bio,gen} \tag{C 4}$$

$$F_{j,t}^{bio,MeOH} = \pi_t \sum_{k\in K} B_{k,j}^{bio,MeOH} \tag{C 5}$$

Generation Technologies

The electricity generated $F_{e,j,t}$ for location j with technology e at time t is generated is proportional to the capacity factor $cf_{e,j,t}$ and generation capacity $Cap_{e,j}^{e}$. Δt is added for unit consistency to convert capacity's MW into MWh for electricity flow in Equation (C 7).

$$\sum_{(k,j)\in A^{bio,MeOH}} B_{k,j,m}^{bio,MeOH} + \sum_{e\in P^{th}} \sum_{(k,j)\in A^{bio,gen}} B_{k,e,j,m}^{bio,gen} + \sum_{(k,l)\in A^{bio,fac}} B_{k,l,m}^{bio,fac} = B_{k,m} \quad \forall k,m \tag{C 6}$$

$$F_{e,j,t} \leq y_{e,j}^{e} cf_{e,j,t}\, Cap_{e,j}^{e}\, \Delta t \tag{C 7}$$

For existing generation, $y_{e,j}^{e}$ is a binary parameter with value of 1.

$$y_{e,j}^{e} = 1 \quad \forall (e,j) \in \varepsilon^{exist} \tag{C 8}$$

For planned generation, $y_{e,j}^{e}$ is a binary variable with value of {0, 1}.

$$y_{e,j}^{e} \in \{0,1\} \quad \forall (e,j) \in \varepsilon^{planned} \tag{C 9}$$

Thus, binary decision variables are used to link capacity investments for both existing and planned assets are defined over $\varepsilon = \varepsilon^{exist} \cup \varepsilon^{planned}$, such that:

$$\underline{Cap^{e}_{e,j}} y^{e}_{e,j} \leq Cap^{e}_{e,j} \leq \overline{Cap^{e}_{e,j}} y^{e}_{e,j} \qquad \forall (e,j) \in \varepsilon \tag{C 10}$$

As thermal power plants are fuelled by biomass resources, additional expression has to be included for the electricity generation flow $F_{e,j,t}$. The generation type involved are only sites running on thermal $e \in P^{th}$.

$$\eta^{e} F^{bio,gen}_{j,t} \geq \sum_{e \in P^{th}} F_{e,j,t} \qquad \forall j \in J, t \in T \tag{C 11}$$

Equation (C 12) can be expressed as the electricity balance at generation, where the overall site balance equation includes term including electricity curtailment $H_{j,t}$, electricity used for methanol production $F^{e,MeOH}_{j,t}$ and electricity generated to be sent to the grid $F^{gen}_{j,t}$. For the equation, terms for storage $Ch^{st}_{e,j,t}$ and discharge $Dis^{st}_{e,j,t}$ are included for the case study, but only exclusively for certain specific hydro dam locations.

$$\sum_{e \in P} F_{e,j,t} + \sum_{e \in P^{hyd}} Dis^{st}_{e,j,t} = F^{gen}_{j,t} + F^{curtail}_{j,t} + \sum_{e \in P^{hyd}} Ch^{st}_{e,j,t} + F^{e,MeOH}_{j,t} \quad \forall j \in J, t \in T \tag{C 12}$$

POC Level

At the POC level, the electricity generated at each generation node is then transported into the transmission network through its connected POC in Equation (C 13). This relationship is represented through the electricity flow variables linking generation sites to POCs. The equation is expressed as:

$$F^{gen}_{j,t} = \sum_{(j,i) \in A^{gen,POC}} F^{e}_{j,i,t} \qquad \forall j \in J, t \in T \tag{C 13}$$

Electricity is then transported across the transmission grid through flows between connected POCs. The set of possible POC connections is defined as $\mathrm{A}^{gen,POC}$, which is a set representing all connections within the transmission network. The expression is illustrated in Equation (C 14).

$$F^{POC}_{i,u,t} \geq 0 \qquad \forall (i,u) \in \mathrm{A}^{POC}, t \in T \tag{C 14}$$

Grid Reinforcement Model

In this section, the constraints that are relevant to grid reinforcement modelling are discussed. These constraints are critically important in the second-stage optimisation framework, as they define the physical and operational limits of the electricity network. The formulations presented in this chapter address several key grid infrastructure components, such as transmission lines, supply transformers, demand transformers, sub-transmission links, and distribution network layers. The key constraints are presented according to the physical hierarchy of the electricity network, beginning with transmission corridors, followed by transformer capacity limits, sub-transmission links, distribution lines and generation expansion.

Transmission Line

Electricity flows between POCs are denoted along transmission layer which links POC nodes i to POC node u. For each connection $(i,u) \in A^{POC}$ and time period $t \in T$, the power transferred through the line is constrained by the rated capacity.

$$F_{i,u,t}^{POC(2)} \leq \overline{Cap_{i,u}^{POC(2)}} \Delta t_t \qquad \forall (i,u) \in A^{POC}, t \in T \tag{C 15}$$

where Δt_t represents the duration of time period t.

When transmission grid is enforced, the existing line capacity can be supplemented by discrete number of upgrades on the lines. The transmission line type is treated as fixed from Transpower asset database. Therefore, the optimisation does not identify preferred line types but instead a discrete replication of the existing capacity rating. The expression is denoted as:

$$F_{i,u,t}^{POC} \leq \left(\overline{Cap}_{i,u}^{POC,0} + \overline{Cap}_{v(i,u)}^{POC,add} n_{i,u,r}^{line}\right) \Delta t_t \quad \forall (i,u) \in A^{POC}, t \in T \tag{C 16}$$

where $v(i,u)$ is the known line type on that link.

Due to the characteristics of transmission connections being bidirectional, the capacity constraint in both directions which are shown in Equation (C 17) and (C 18) are the same for each connected pair $\{i,u\}$.

$$F_{i,u,t}^{POC} \leq Cap_{i,u}^{POC} \Delta t_t \tag{C 17}$$

$$F_{u,i,t}^{POC} \leq Cap_{i,u}^{POC} \Delta t_t \tag{C 18}$$

$$\forall t \in T, \forall \{i,u\} \in A^{POC} \tag{C 19}$$

For the integer sizing approach implemented using the P-graph framework, a maximum allowable number of grid upgrades must be assigned within the model. This constraint can be represented mathematically as:

$$0 \leq n_{i,u}^{line} \leq n_{i,u}^{line,max}, \qquad n_{i,u}^{line} \in \mathbb{Z} \qquad \forall (i,u) \in A^{POC} \tag{C 20}$$

This limit is important because, in the P-graph representation, each discrete upgrade unit is modelled as a standalone structural node. As a result, the number of allowable upgrades must be bounded explicitly to control the size of the superstructure and maintain computational tractability.

The cost expression for the integer sizing or number of grid upgrades is represented by Equation (C 21). The cost coefficient $C_{(i,u)\in A^{POC}}^{trans,line}$ reflects the investment cost associated with reinforcing the transmission corridor between two connected POCs. This coefficient depends on the voltage rating of the line type and the physical length of the link.

$$CC^{trans,line} = \sum_{(i,u)\in A^{POC}} C_{(i,u)\in A^{POC}}^{trans,line} n_{i,u}^{trans,line} \tag{C 21}$$

Supply Transformer

The supply side transformer capacity at each POC restricts the total electricity flow from generation sites. The inflow of electricity is defined as Equation (C 22):

$$F_{i,t}^{POC,in} = \sum_{j \in J} F_{j,i,t} \quad \forall i \in I, t \in T \tag{C 22}$$

Transformer component is represented in Equation (C 23) using a discrete sizing structure, where multiple standard transformer capacity blocks may be installed. The inflow constraint is therefore written as:

$$F_{i,t}^{POC,in} \leq \left(\sum_{s \in S^{tr}} \overline{P_s^{tr}} n_{i,s}^{POC,in} \right) \Delta t_t \quad \forall i \in I, t \in T \tag{C 23}$$

where each element $s \in S^{tr}$ acts as a serial selector which bundles the rated capacity and discrete reinforcement step, and $\overline{P_s^{tr}}$ denotes the corresponding rated capacity. The integer variables $n_{i,s}^{POC,in}$ determines the selection index, not a count constraint.

The total supply transformer reinforcement cost is represented by:

$$CC^{tr,sup} = \sum_{i \in I} \sum_{s \in S^{tr}} C_s^{tr} \, n_{i,s}^{POC,in} \tag{C 24}$$

Demand Transformer

The same theory applies to demand transformer by where electricity is leaving the POC node to be distributed to the demand for the corresponding node as well as distribution to the industrial sites. In Equation (C 25), the outflow of electricity from a POC in relation to the discrete sizing structure of the transformer component can be described as:

$$F_{i,t}^{POC,out} \leq \left(\sum_{s \in S^{tr}} \overline{P_s^{tr}} n_{i,s}^{POC,out} \right) \Delta t_t \quad \forall i \in I, t \in T \tag{C 25}$$

The similar cost function applies for the demand transformer reinforcement where it varies according to the capacity size.

$$CC^{tr,dem} = \sum_{i \in I} \sum_{s \in S^{tr}} C_s^{tr} \, n_{i,s}^{POC,out} \tag{C 26}$$

Sub transmission and Distribution Line

As for the distribution and sub-transmission layers, grid reinforcement is represented through discrete line-type selection. Instead of allowing continuous expansion, typically each connection is only assigned to a single capacity option from a predefined set of standard line ratings type. Binary decision variables are determined for the given line type based on the electricity load carried in the line. The selected rating can then be defined as the line upgrades for new generation sites for sub-transmission layer and for large process heat use for distribution layer.

Sub transmission line

Electricity flows from generation nodes $j \in J$ into the transmission grid at POC $i \in I$ through the sub-transmission layer.

$$F_{j,i,t}^{e} \leq \sum_{r \in R^{subtrans}} \left(\overline{Cap}_{j,ir}^{subtrans} \, y_{j,i,r}^{subtrans} \right) \Delta t_t \quad \forall (j,i) \in A^{gen}, t \in T \tag{C 27}$$

where $r \in R^{subtrans}$ indexes the available line capacity types. $\overline{Cap}_{j,ir}^{subtrans}$ is the rated capacity of line type r and $y_{j,i,r}^{subtrans}$ is the binary decision for installation of the specified line. Rule is set for line selection to be only a single line type can be expressed as:

$$\sum_{r \in R^{subtrans}} y_{j,i,r}^{subtrans} \leq 1 \qquad \forall (j,i) \in A^{gen} \tag{C 28}$$

$$y_{j,i,r}^{subtrans} \in \{0,1\} \quad \forall (j,i) \in A^{gen}, r \in R^{subtrans} \tag{C 29}$$

The total sub-transmission investment cost of is presented as:

$$CC^{subtrans} = \sum_{(j,i) \in A^{gen}} \sum_{r \in R^{subtrans}} C_{j,i,r}^{subtrans} y_{j,i,r}^{subtrans} \tag{C 30}$$

where $C_{j,i,r}^{subtrans}$ represents the cost of installing line type r along the connection (j,i).

Distribution lines

Similarly, the electricity delivered from each POC node i to industrial demand $l \in L$ is denoted as constrained through distribution line capacity sizing:

$$F_{i,l,t} \leq \sum_{r \in R^d} \left(\overline{Cap}_{i,l,r}^{dist} y_{i,l,r}^{dist} \right) \Delta t_t \tag{C 31}$$

As both sub-transmission line and distribution line share the same modular operations, same rule applies for distribution line to choose only a single line per connection. Equation (C 32) enforces the constraint such that:

$$\sum_{r \in R^{distribution}} y_{i,l,r}^{dist} \leq 1 \qquad \forall (i,l) \in A^{dem} \tag{C 32}$$

The distribution connection is a binary constraint enforcement:

$$y_{i,l,r}^{dist} \in \{0,1\} \qquad \forall (i,l) \in A^{POC,fac}, r \in R^d \tag{C 33}$$

The investment cost incurred for each line connection is presented in Equation (C 34) as:

$$CC^{dist} = \sum_{(i,l) \in A^{dem}} \sum_{r \in R^d} C_{i,l,r}^{dist} y_{i,l,r}^{dist} \tag{C 34}$$

Generation Component

The equations for accounting the additional electricity generation sites for accommodating additional electricity requirement are presented as below. The electricity generation upgrade cost for new generations will be added into the total cost for the minimisation of the objective function.

$$E_{e,j,t} \leq y_{e,j}^{e} cf_{e,j,t} \, Cap_{e,j}^{e} \, \Delta t \tag{C 35}$$

For planned generation, $y_{e,j}^{e}$ is a binary variable with value of {0, 1}.

$$y_{e,j}^{e} \in \{0,1\} \quad \forall (e,j) \in \varepsilon^{planned} \tag{C 36}$$

The total investment cost for generation expansion is then expressed as:

$$CC^{gen} = \sum_{e \in E} \sum_{j \in J} C_{e,j}^{gen} Cap_{e,j}^{e} \quad \text{(C 37)}$$

where $C_{e,j}^{gen}$ denotes the unit capital cost of installing generation technology e at location j.

Appendix D

Appendix D describes the grid upgrade costs used in the model. obtained from Ergo Consulting (2024) during the preparation of the Regional Energy Transition Accelerator reports on behalf of EECA. Table D 1 outlines the costs for transmission equipment, Table D 2 presents the sub-transmission equipment costs, whereas Table D 3 shows the distribution equipment costs.

Table D 1: Transmission equipment cost assumptions used in the grid reinforcement model.

Equipment	Quantity	Operating Range	Cost (NZD m)
220kV single circuit line	km		1
220kV double circuit line	km		2
Large GXP	unit		20
Large supply transformer (GXP)	unit	100-200 MVA	9
Medium supply transformer (GXP)	unit	50-100 MVA	7
110kV circuit breaker bay	unit		1.2
110kV single circuit line	km	50-100 MVA*	0.8
Reconductoring of existing 110kV line	km		0.6
Single overhead 66kV line	km		0.35
Double overhead 66kV line	km		0.6
Special protection scheme (GXP)	unit		0.5
110kV double circuit line	km		1.8

Table D 2: Sub transmission equipment cost assumptions used in the grid reinforcement model.

Equipment	Quantity	Operating Range	Cost (NZD m)
Single underground 33kV cable	km		0.9
Single overhead 33kV line	km	12 MVA*	0.35
Double overhead 33kV line	km		0.35
Double underground 33kV cable	km		1
33kV circuit breaker bay	unit		0.25
Medium switchroom (Zone Substation - ZSS)	unit		6
Medium supply transformer (ZS)	unit	15-30 MVA	1.9
Large supply transformer (ZS)	unit	30-60 MVA	2.3
Reconductoring of existing 33kV line	km		0.45

Table D 3: Distribution equipment cost assumptions used in the grid reinforcement model.

Equipment	Quantity	Operating Range	Cost (NZD m)
11kV circuit breaker (ZSS)	unit		0.1
33kV circuit breaker (ZSS)	unit		0.15
Single underground 11kV cable	km		0.6
Single overhead 11kV cable	km	4.5 MVA*	0.2
Small zone substation (ZSS)	unit		8
Medium zone substation (ZSS)	unit		15
Double underground 33kV cable	km		1
200kVA distribution transformer	unit		0.08
750kVA distribution transformer	unit		0.2
Medium supply transformer (ZS)	unit		1.9
Distribution switches - RMU	unit		0.05
Reconductor 11kV line	km		0.15
11kV capacitor bank	unit		0.3

Appendix E

Source Codes Included in the Thesis Repository

This appendix only contains key source codes developed for the thesis; supporting spreadsheets, input data files, and other external resources are excluded.

To preserve the confidentiality of the data. Data values have been removed and the directory for file paths have been generalised. The GitHub repository is as follows:

https://github.com/danieljschong/Process-Heat-Energy-Modelling-Framework

E1 P-graph Library Modification

Bugs on operating unit nodes are fixed and nodes with additional non-solving attributes such as units are added for better data handling for the data output.

E2 Porting and Solving Tool

This section contains the P-graph-to-Pyomo translation tool together with the regular-expression rules used to convert the superstructure into an optimisation model.

E.3 Preprocessing Tools

Renewables Ninja - This subsection contains the Python-based scripts used to access weather and intermittent generation data through the Renewables Ninja API.

ArcPy Tools – Distance Accumulation Tools

ChromeDriver – Electricity Spot Price and Demand for each trading period

E.4 Output Processing

This section contains the output-processing scripts used for masking, formatting, and presenting model results in tabular form.

E.5 Simulation Modes

This section contains the source codes implementing the main simulation modes used in the thesis.

Multiperiod Deterministic Optimisation – contains base model used in Chapter 7

Sequential Optimisation with Root-Finding - contains the code for the sequential optimisation approach coupled with root-finding procedures.

Monte Carlo Uncertainty Analysis - This subsection contains the code for the Monte Carlo-based uncertainty analysis.

Appendix F

F1.1 Generation

The electricity generation is an important component to assess the process heat fuel-switching decarbonisation in the New Zealand industry. Figure F 1 illustrates the modular structure of a generic commissioned electricity generation site. A raw material node with a prefix of 16 is introduced into the model and connected to an operating unit node that determines the generation output based on the installed capacity of the site. In principle, the raw material node does not require an upper boundary because the generation limits are already constrained by the installed capacity of the associated unit, identified by the prefix 14.

In practice, however, any custom upper boundary applied to this raw material node is later replaced when biomass raw material nodes are introduced. These biomass nodes represent forest-based resources that supply thermal, combined heat and power, and bioenergy generation sites. This substitution sets the physical availability limit for biomass and reflects competition between electricity generation, process heat users, and emerging green industry consumers.

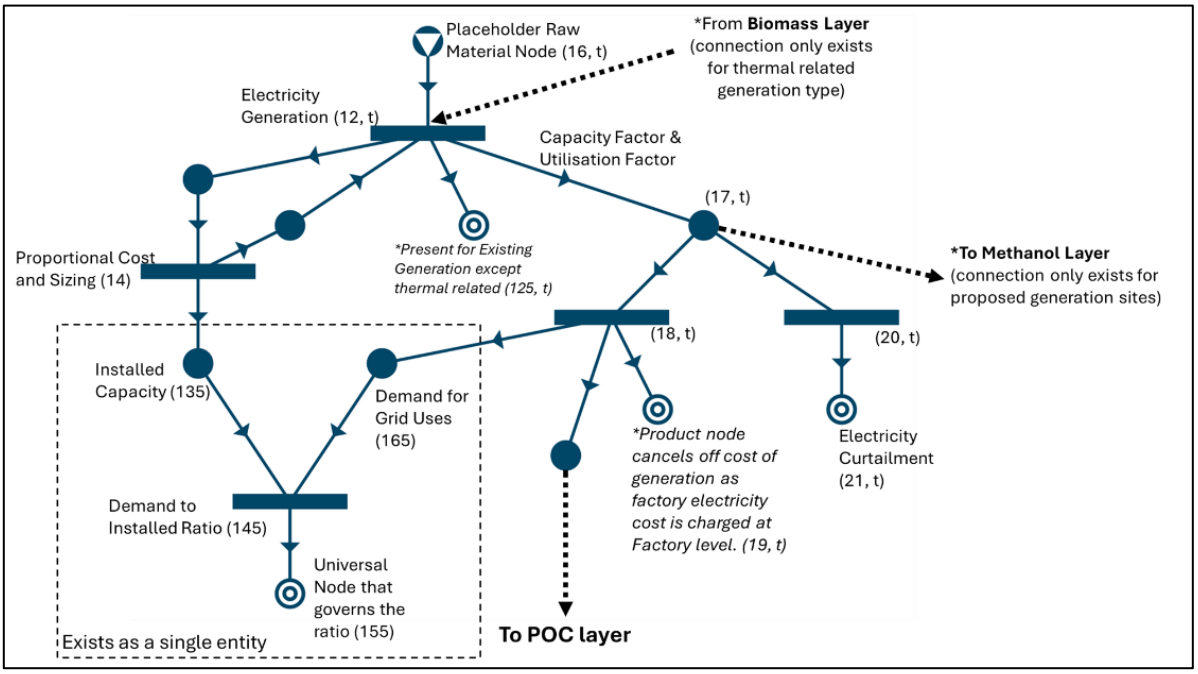


Figure F 1: Modular component for the generation layer with |T|=1.

As shown in Figure F 1, the operating node with an ID prefixed by 12 contains a time series element. This arises because the weight factor linking the operating node to the material node with a prefix of 17 is the capacity factor for the given time period. Under a monthly time step, the weight factor represents the total available generation in that month, calculated as the monthly capacity factor multiplied by the number of hours in the month. Under an hourly time step, the weight factor corresponds directly to the capacity factor for that hour.

Three outward edges extend from the material node with a prefix of 17. One of these leads to an operating unit node with a prefix of 18, which represents the pathway to the electricity grid. One edge connects to the green methanol site, while additional edges link to node with prefix of 21 serves as a placeholder for electricity curtailment when generation is in excess. The operating unit node with a prefix of 18 represents the sub transmission charges incurred when a generation unit is connected to the grid. As a result, a submodule that represents the installed capacity of the sub transmission delivery lines is added to account for the capital costs of new network installations.

The product node associated with the operating unit with a prefix of 18 performs a balancing calculation that offsets the electricity cost from both existing and new generation against the installed capacity cost. This reflects how electricity prices paid by consumers are determined through negotiations with gentailers under fixed-price variable-volume (FPVV) arrangements. Retaining the installed capacity cost within node prefix 14 is beneficial for the techno-economic feasibility assessment of green methanol production sites. The operating unit with prefix 18 subsequently connects to the POC node.

The universal node with an ID of 155 compares the electricity flow to the grid with the installed capacity. This provides a graphical method to enforce a ratio of electricity use to installed capacity of 0.6 or lower, maintaining electricity security when the model allows grid supplied electricity to meet process heat demand. In the PNS formulation, this goal is achieved by applying a multiplier of 0.6 along the edges connecting the nodes with prefixes 14 and 135. Furthermore, setting the upper boundary of the node with prefix 135 to zero, while allowing the node with prefix 165 to an unbounded upper limit, ensures that electricity use across all installed capacity remains at or below the 0.6 threshold. Surplus electricity flows can accumulate at the node with prefix 165, serving as a boundary relaxation mechanism to preserve feasibility in the optimisation.

The node with the prefix 125 is included only for commissioned plants to ensure that the relevant substructure is always present in the final solution. Its lower boundary is set equal to the installed capacity of the generation site, consistent with Axiom 1 of the P-graph framework. The addition of this product node reduces the computational burden on the solver because the generation component is enforced to appear in the solution, thereby avoiding unnecessary discrete decisions.

F1.2 Biomass

The biomass modular structure is constructed differently from the other components, as it is only activated when specific criteria are met. At a broader level, satisfying the general model settings, such as geographical conditions and allowable travel distances, is not sufficient. Each spatially defined biomass node must also appear in at least one of the three possible pathways: (1) biomass supplied to factories, (2) biomass supplied to generation sites, or (3) biomass supplied to green methanol plants. This requirement creates a prerequisite in which the biomass structure can only be built when the associated destination modules exist within the model. In effect, the biomass database functions as both a standalone representation of resource availability and location and as a network of interactions that link each biomass source to its eligible destinations using the preprocessing methodology mentioned in Section B.

Biomass Joining to Factories

The biomass available resource node has an identifier prefix number of 8 while the suffix represents the resource type in the location. In the framework, 5 biomass types m, denoted as {0, 1, 2, 3, 4}, are deployed. The biomass-to-factory connection is established with a unique ID combination where the biomass identifier is appended with the factory identifier, creating a unique identifier. The operating nodes which are shown in Figure F 2, which connect the biomass resource nodes, account for biomass transport costs, with the associated distinct distance values represented in the cost function of the operating node. It is not only unique for each transport connection but also among different biomass resource types, as it contains different energy densities.

The material node with a suffix of 150 serves in a cumulative manner to total the amount of useful biomass which is categorised based on the combustion technology they are passing through. The node structure must be established to manage the large number of possible transport routes, totalling 27,563 (or 330,756 when accounting for a 12-time step multiperiod model). The operating units accounting for distance costs are then joined and grouped into two distinct nodes, 3000+j+P^{th}, where P^{th} = {biomass boiler: 0, bubbling fluidised bed: 1}. The node with a base identifier of 140+P^{th} and a time dimension t is joined to a substructure operation which accounts the proportional sizing for both boiler capacities and heat energy supplied for different time steps. Note that the stream exiting the substructure accounts for the thermal efficiency and energy transformation of the biomass feedstock before branching into the factory demand node with the identifier of 124+a, where a accounts for the temperature quality in ascending order.

An alternative retrofitted biomass boiler from coal with the base identifier of 160 and a time dimension t provides a retrofitting option for factory users where the capacity is predefined and competes with other heating technologies in supplying heat in the P-graph structure.

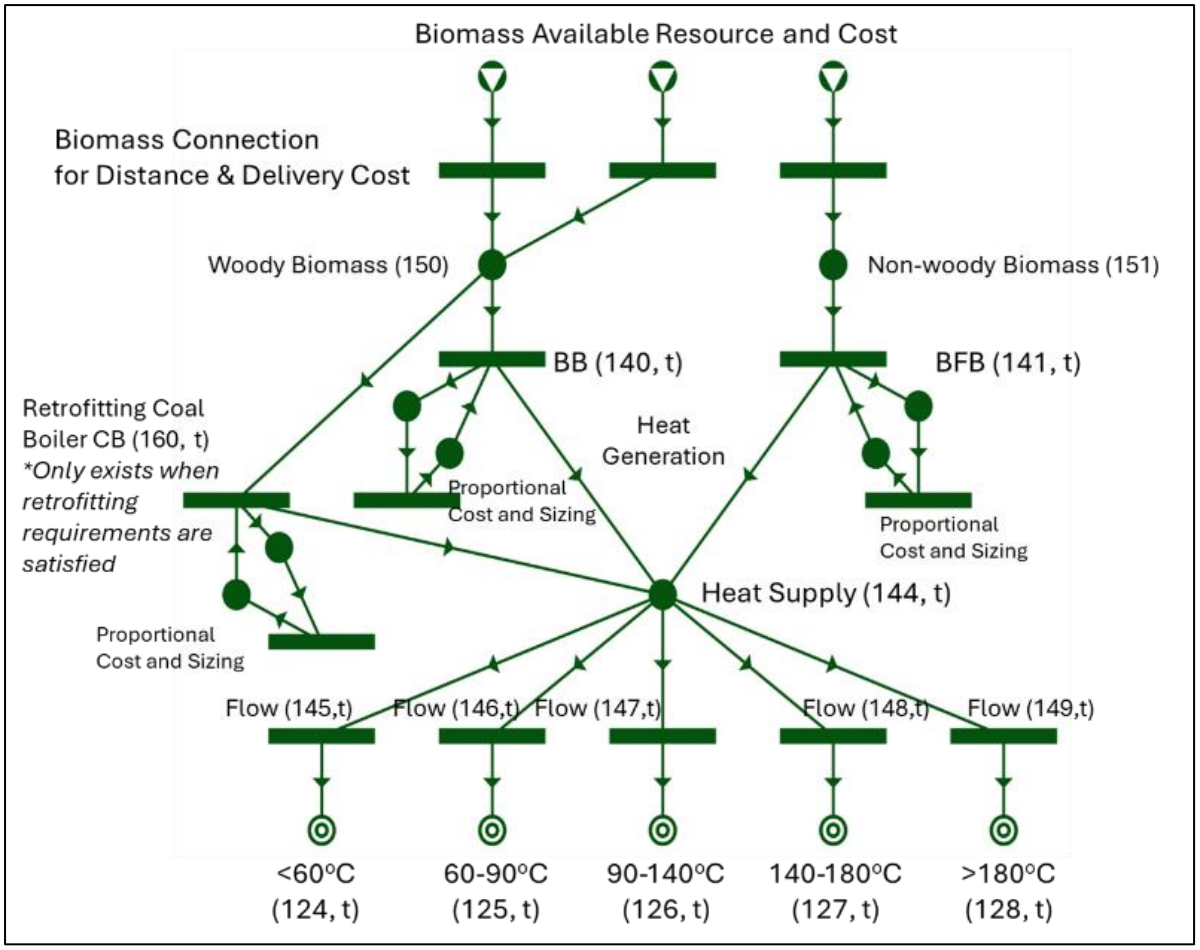


Figure F 2: Modular component for the biomass to factory layer with |T|=1.

Biomass Joining to Methanol Plants and Generation Plants

The connection database for methanol plants and generation plants is shared, because methanol production facilities are co-located only at electricity generation sites that have not yet been commissioned. Biomass is used as the primary fuel for thermal-based generation sites that operate steam turbine units. As shown in Figure F 3, each green methanol plant is linked to a biomass supply node where the biomass ID is appended with the corresponding generation site ID. The operating unit associated with this link applies the cost function that accounts for distance-based transport costs and the conversion of energy units from gigajoules to megawatt hours. This operating unit also outputs the required quantity of biomass specific to each resource type.

The node identified by a suffix of 170 aggregates the biomass flows entering the methanol plant. The inward operating unit node with a middle string of 250 contains coefficients on its edges that represent the conversion requirements of electricity and biomass needed to produce one tonne of green methanol. A sub-modular structure is then introduced to determine the minimum system cost of green methanol production by assembling all relevant substructures across the full time series. The cost in the operating unit node with a suffix of 170 represents the capital and fixed operating costs, while the variable operating costs arise from the biomass resource and transport as well as the electricity generation site, which both provide the feedstock required for

green methanol production. The product node with a middle identifier of 260, appended with the corresponding time period, reports the quantity of green methanol produced at each time step.

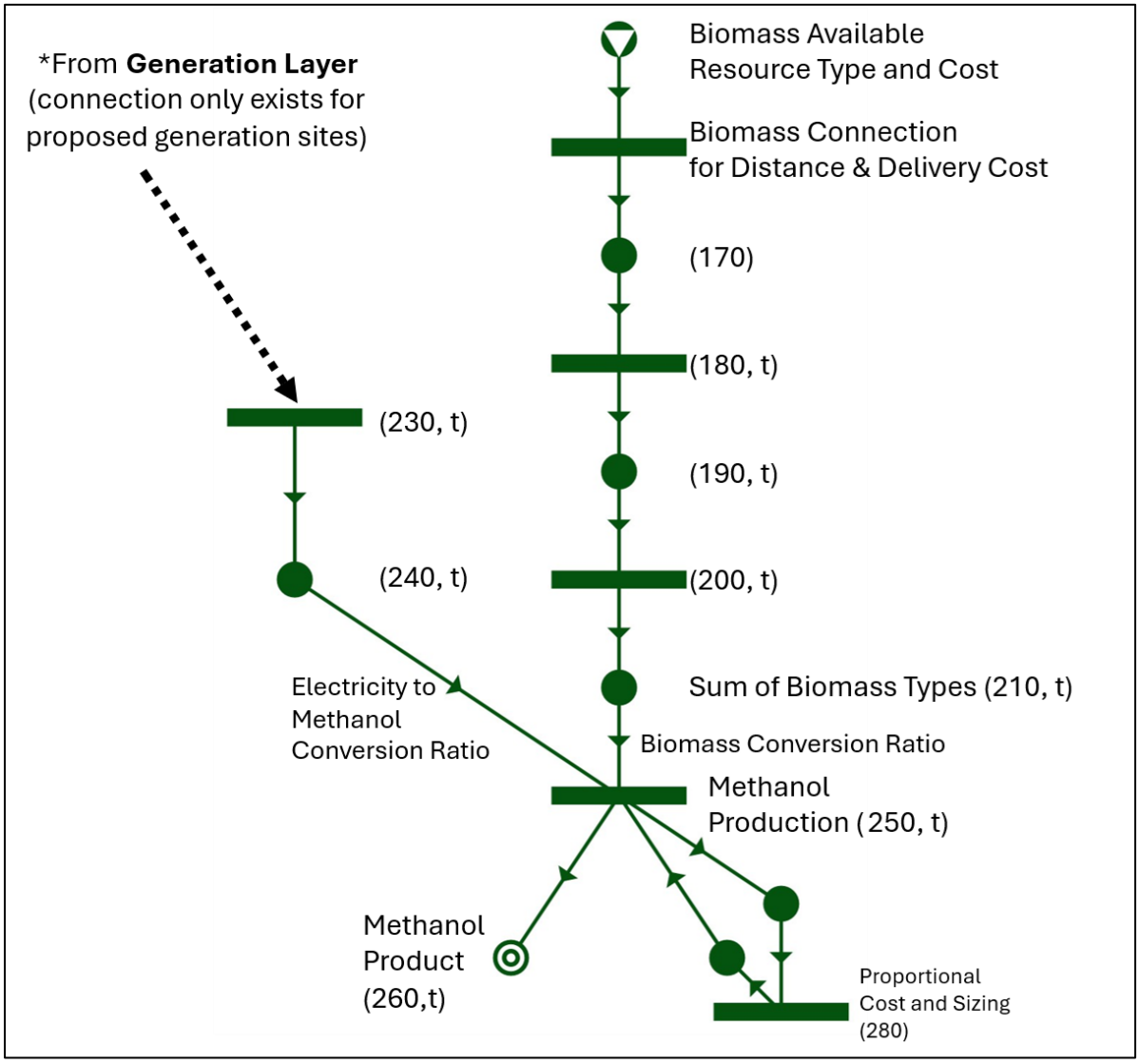


Figure F 3: Modular component for the biomass to methanol layer with |T|=1.

On the biomass to electricity generation side, the operating unit node associated with each biomass resource is assigned with an identifier that appends the biomass ID to the corresponding generation site ID, with a middle identifier of 300. This is to distinguish the biomass-generation pathway from other transport uses. Biomass fuels from all eligible locations are first aggregated and then connected to the operating unit node, which is prefixed by 12 and contains a time series component. This operating unit represents the energy transformation process for electricity generation, with the time-dependent element capturing variation in generation output across the selected temporal resolution. A product node with a middle identifier of 400 is in place to eliminate additional charges to the factory layer model on biomass-powered electricity generation cost, as electricity price has already been charged at the factory level. The modular structure of this connection is shown in Figure F 4.

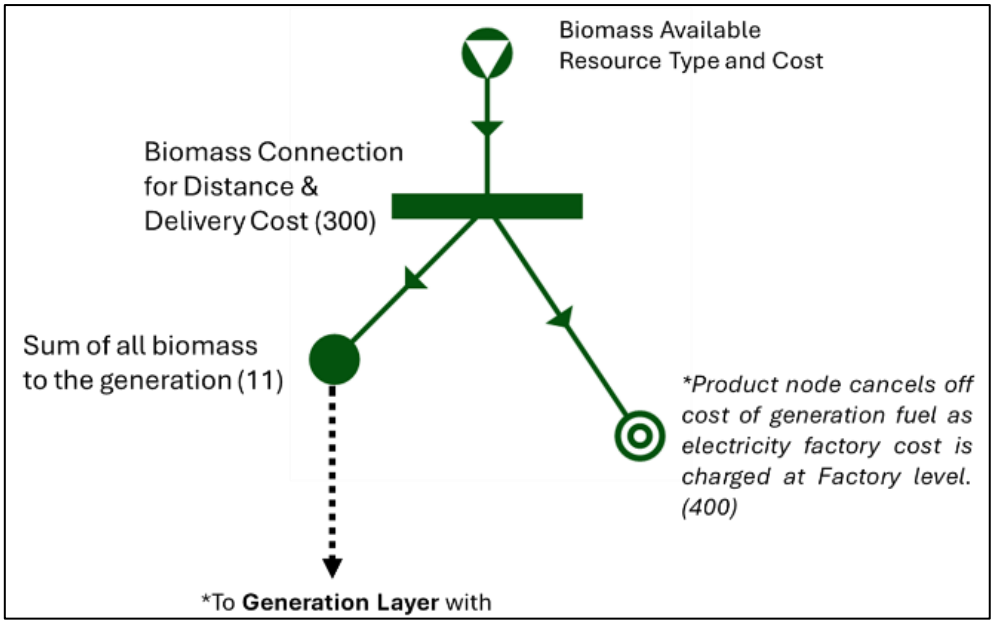


Figure F 4: Modular component for the biomass to generation layer with |T|=1.

F1.3 Hydro-storage

The hydro storage modular structure is incorporated into the model only when the conditions specified in the generation prompts are met. The generation type must be hydroelectricity with a dam, since run-of-river schemes do not provide medium- or long-term storage capability. Moreover, the storage representation is limited to the immediate upstream water hold up entering the power station, which reduces the overall model scale and allows greater analytical detail to be directed toward process heat demand in industrial plants. The focus of the model is to evaluate site-level process heat decarbonisation where full hydro dispatch optimisation though it affects the electricity dynamics, is not the scope of the model.

Besides that, the spill events or flood control operations have to be studied for full hydro because the optimisation does not follow the minimum cost-oriented optimisation problem. Estimations on hydrological inflow and outflow uncertainty would also be required but would exacerbate the proper analysis of the energy model with a large uncertainty variable. As a result, the process network is streamlined to follow the sequence from generation to transmission and then to factories, rather than modelling the full mesh of interconnected streams that link multiple hydroelectric stations within the same hydro scheme. The simplified approach is adequate because hydro is treated as a stable baseline rather than a system margin resource, hence the approach is sufficient to evaluate the influence of hydro storage on the electrification pathways of the heating technologies in the factories.

Referring to the P-graph superstructure representing the modular operations of hydro storage, the material node from the Generation Layer can be branched off to an operating unit node ID with a prefix of 16. The operating unit possesses a time series which is important to show that additional electricity, which is not consumed, can be stored for the subsequent time step. The

hold-up node handling the inventory of the lake storage is handled by the same operating unit ID. The operating ID with prefix 16 and suffix 3 is the main constraint for contingency water storage and maximum lake water can be stored before expending it for electricity generation. For simplification, the quantification unit for the water storage component is in GWh instead of the volume of stored water as it is easier for unit conversion for drawing and storing electricity from the lakes for carryover between time steps. It is also particularly critical, as each hydroelectricity plant factor is different across all plants in terms of electricity generated from water throughput. Figure F 5 shows the extended substructure implemented for hydroelectricity generation sites with hydro dams.

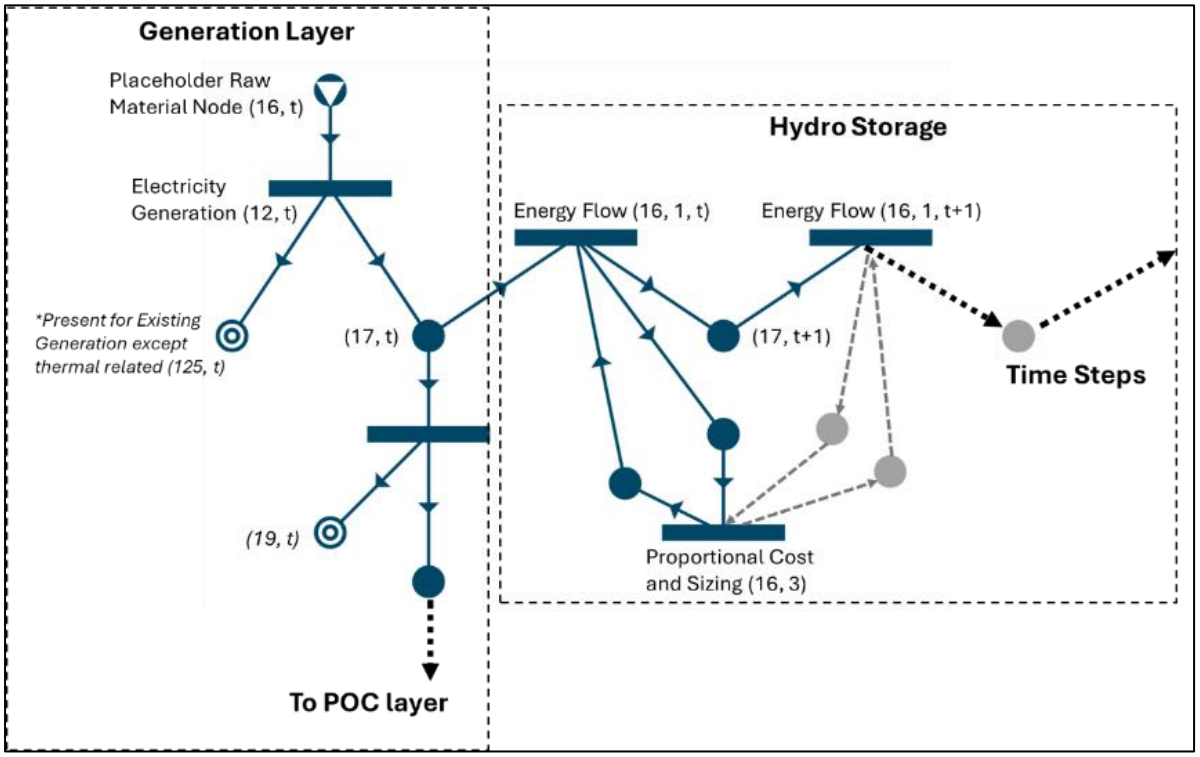


Figure F 5: Modular component for the hydro storage extension at the generation layer with |T|=1.

Similar to the same concept for equation calculations for cyclic boundary condition, the last time step has to loop back to the first-time step, so as to allow accurate representation for a steady-state horizon problem. The loop also prevents bias created by a fixed initial storage level as lake level fluctuates due to inflow, direct rainfall, groundwater recharge, spill, seepage and other related outflows.

$$S_{t+1} = S_t + \eta_{ch} I_t - \frac{1}{\eta_{dis}} O_t^w \quad \forall t = 1,2,\dots,T-1 \qquad (0.1)$$

The cyclic boundary condition is imposed as:

$$S_1 = S_T + \eta_{ch} I_T - \frac{1}{\eta_{dis}} O_T^w \qquad (0.2)$$

where S_t is the storage level at time step t, I_t denotes the inflow or charging at time t, O_t^w is the outflow or discharging at time t, η_{ch} is the charging efficiency, η_{dis} is the discharging efficiency and T is the total number of time steps in the horizon.

F1.4 Point of Connections (POCs)

Grid modelling

The modular framework for the POC is constructed using two interconnected data structures: a single-dimensional database and a multidimensional database. The single-dimensional layer defines the fundamental attributes and identifiers required for each POC, while the multidimensional layer captures the additional spatial interactions and technical-related characteristics needed to represent the full complexity of the inter-POC transmission grid system.

The single-dimensional database contains the essential attributes associated with each POC. These include the POC name, its unique identifier, annual electricity demand by year, wholesale electricity price profiles, expected process heat demand, N security demand, N security generation, geographic coordinates, and other general regional information.

In contrast, the multidimensional database represents the network relationships between POCs. It includes attributes such as the source and destination POCs, the distance between them, the island or regional classification, grid identifiers, grid asset capabilities, transmission voltage levels, and the corresponding line types.

Figure F 6 illustrates the generic modular structure of the POC layer. Using the single-dimensional database, the streamlined flow from the material node with a middle identifier of 101 through to the product node with the prefix 43 can be constructed. The material node with identifier 101 represents the N security requirement for generation, which is simplified in this model to the transformer capability required to step up electricity to transmission voltage levels.

Likewise, the operating unit node with the prefix 23 represents the N security requirement for demand at the POC level. This node branches into two separate categories: one that connects to the factory-level modular structure, and another that reflects non-industrial, non-process heat-related electricity consumption within the POC. The product node includes residential and commercial electricity use, electric vehicle charging, and non-process industrial electricity demand, such as lighting, stationary equipment, motors, fans, and other ancillary loads.

The second category connected at the distribution level is the factory node, where the number of additional branches ranges from zero to the total number of factories associated with a given POC. The POC module and the factory module are linked through a unique factory identifier, since the factory database contains explicit information on which POC site is attached to. Each site-level operating unit is assigned a distinct cost function that reflects the electricity prices faced by industrial users for additional process heat demand. These operating unit costs also incorporate network-related charges, which depend on both installed capacity and the distance between the site and the relevant grid assets. Because these network cost components vary across locations, each factory node in the New Zealand case study has a unique associated cost.

Assigning cost coefficients that scale with the energy flow further enhances the performance of the optimisation. This proportional cost structure strengthens the guidance provided to the branch and bound solver by allowing it to prioritise subproblems with lower marginal costs. As a result, the solver can focus its search on subproblems that are more likely to yield feasible solutions contributing to the global minimum of the superstructure, thereby reducing computational effort and improving solution efficiency.

At the transmission level, a simplified energy flow superstructure is constructed to represent the connections among all POCs. This transmission layer can be formulated in two distinct versions. The first version applies no boundary constraints or uses only broad, arbitrary limits defined as multiples of the existing grid capacity, which allows for sensibility checks and preliminary exploration of feasible flow patterns. Such an approach is useful for identifying whether unrealistic or extreme grid utilisation would be required in the absence of upgrades. For instance, imposing an upper bound equal to twenty times the current line capacity would be illogical and therefore serves as an example of a constraint that should not be used.

The detailed version of the transmission structure incorporates the explicit grid upgrade layer. In this formulation, capacity constraints are set in accordance with plausible upgrade options, enabling a more realistic evaluation of network reinforcement requirements. The detailed version will be discussed in the next section.

In the transmission substructure, the material node with the prefix 2 represents the POC, which serves as the common element linking the modular frameworks of both the POC level and the transmission level. The multidimensional database contains the set of POC-to-POC connections, and this information is used to construct the superstructure through which electricity can flow across the network. Each row in this database corresponds to a POC edge, and the index of each entry is translated into the identifier of an operating unit node by assigning an additional prefix of 6, ensuring that these nodes are uniquely distinguished within the system.

The electricity flows are modelled as unidirectional at any given time step, a corresponding return edge is created for each pair of connected POCs. Consequently, the operating unit nodes with identifier prefix 6 that represent the forward and reverse flows between the same two POCs at the same time step must include a disjunctive property, ensuring that only one direction of flow being active in the superstructure at any time. This prevents simultaneous bidirectional flow and maintains the physical consistency of the network representation.

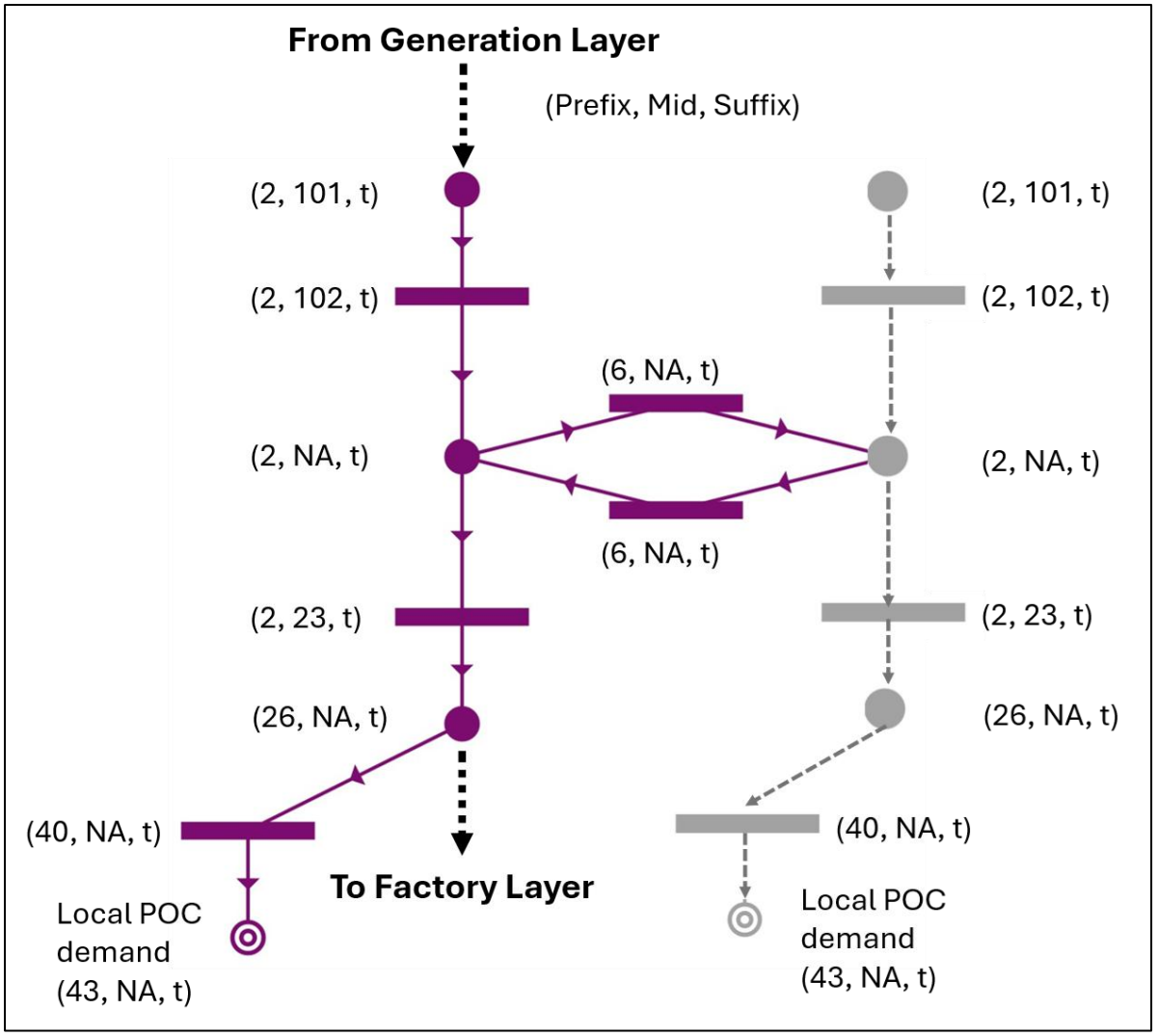


Figure F 6: Modular component for the transmission layer with |T|=1.

F1.5 Advanced Grid Modelling

Additional nodes can be incorporated into the energy model to evaluate the condition and adequacy of electricity grid assets. These ancillary grid upgrade modules include the costs associated with distribution line reinforcements required to supply additional electricity to factories, on-grid charges for electricity generation stations, step-up and step-down transformer upgrades at the POC level in Figure F 7, and transmission capacity enhancements within the National Grid. Although these modular structures are optional, they play an important role in feasibility assessments, particularly when planning future grid infrastructure as an implication of the increased electricity demand arising from industrial process heat electrification.

The structural method employed here follows the same approach presented in the methodology section of the journal article titled “Optimal Sizing of Solar Powered Panels at a Residential Level by P-graph”. The goal is to formulate a discrete decision variable of components in the grid to optimally design technological options with different specifications while considering the

different costings which are considered stepwise costs when procuring and installing electricity components with the specified operating capabilities. The total annualised costs for these components with different specifications are stored in a cost lookup module, which can be imported in Python. The case study of applying grid analysis will be explained in Chapter 8 as part of a two-step optimisation to solve the grid reinforcement model.

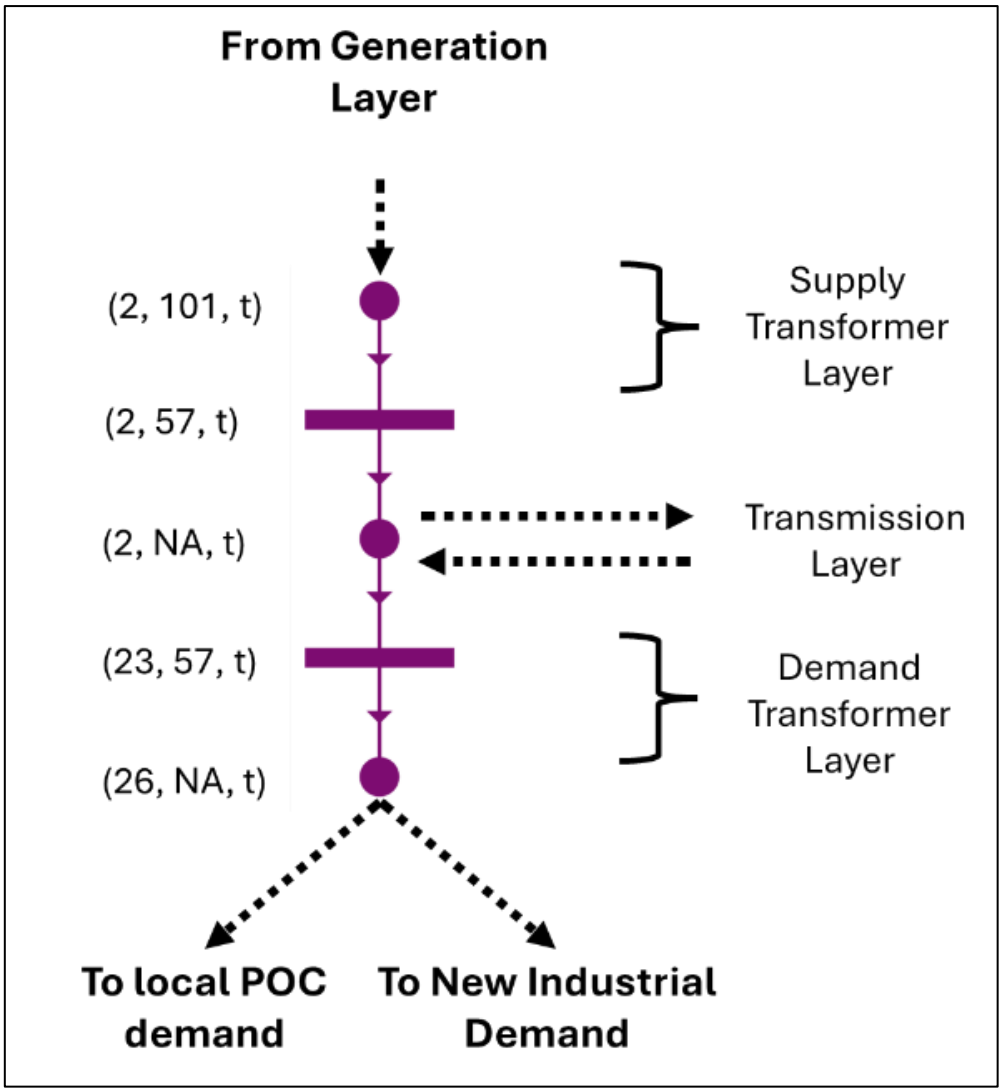


Figure F 7: P-graph structures across different layers of the electrical grid system.

Sub-transmission and Distribution

Both sub-transmission and distribution capacity upgrade lines share the same cost lookup module. Despite having different functionality in the electricity network, as one draws electricity and one supplies electricity from the grid; both upgrade costs are based on the new allowable lines needed to be connected to substations or POCs. From the proposed electricity generation perspective, the sizing of the power station is determined by the optimisation based on the required demand projections. However, the maximum installed capacity is set according to the proposed plans disclosed to the public. Therefore, a series of distribution lines with target specifications is introduced when the modular structure is implemented in the model. Referring to the line

specification table, the lines are defined based on the load they can handle, enabling all six potential lines to exist in the initial problem.

Moreover, the time resolution setting allows a better representation of the energy flow that the lines can withstand, as the line capacities are defined in terms of load (MW). From the cost lookup table, the lower and upper bounds are not set cumulatively based on the load size in ascending order. This is because treating the distribution lines as independent from one another improves boundary relaxation for unsteady energy flow across the grid in certain time steps. In addition, the combinatorial configuration allows the model to assess whether installing two distribution lines could be more cost-effective than installing a single line with a larger load capacity.

Furthermore, the cost of each electricity line is based on the distance between the energy players (here, the generator) and the POC. The line connections are designed to be distinct across all sites, thereby reducing the pruning effect for the algorithms when searching for the global minimum solution. In the model, new connections are only needed for large electricity users, as electricity connection has already been established for powering machinery, lights and motors. A similar concept has been applied with the main difference having to be that the distribution lines would be parallel to the electricity demands for non-industrial heat users, instead of at the generation level.

The operating unit that represents the maximum load capacity shares the same identifying prefix of 18, but includes an index n, where n corresponds to the available line specifications listed in the cost lookup module in ascending order. Figure F 8(a) illustrates the upgrades for sub-transmission lines and Figure F 8(b) presents the upgrades for distribution level.

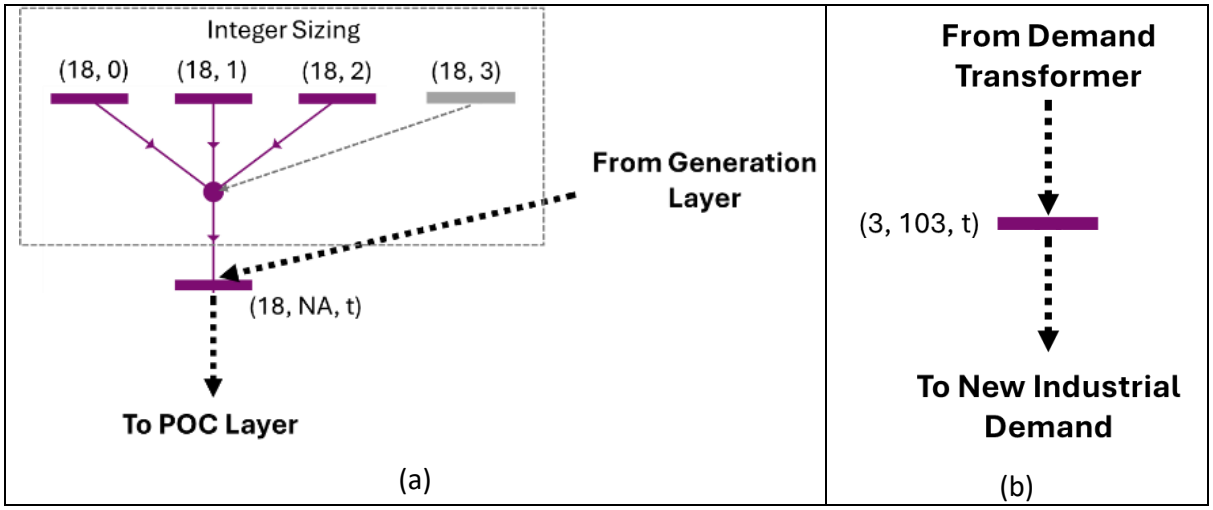


Figure F 8: P-graph representation of sub-transmission lines and distribution lines when |T|=1.

Transmission

The transmission layer of the model represents the amount of existing electricity grid infrastructure required to deliver electricity from various segments of the network to demand centres, such as residential areas and industrial process heat users. Because available information on grid capacity (in MVA) is reported at the voltage level, and attributes such as line distance or line type (for example, single circuit, double circuit, or underground cables) do not translate directly into capacity values, the model assumes that transmission capacity is proportional to the number of existing lines connecting the relevant POCs. The upgrades by integer sizing in the transmission layer are shown in Figure F 9.

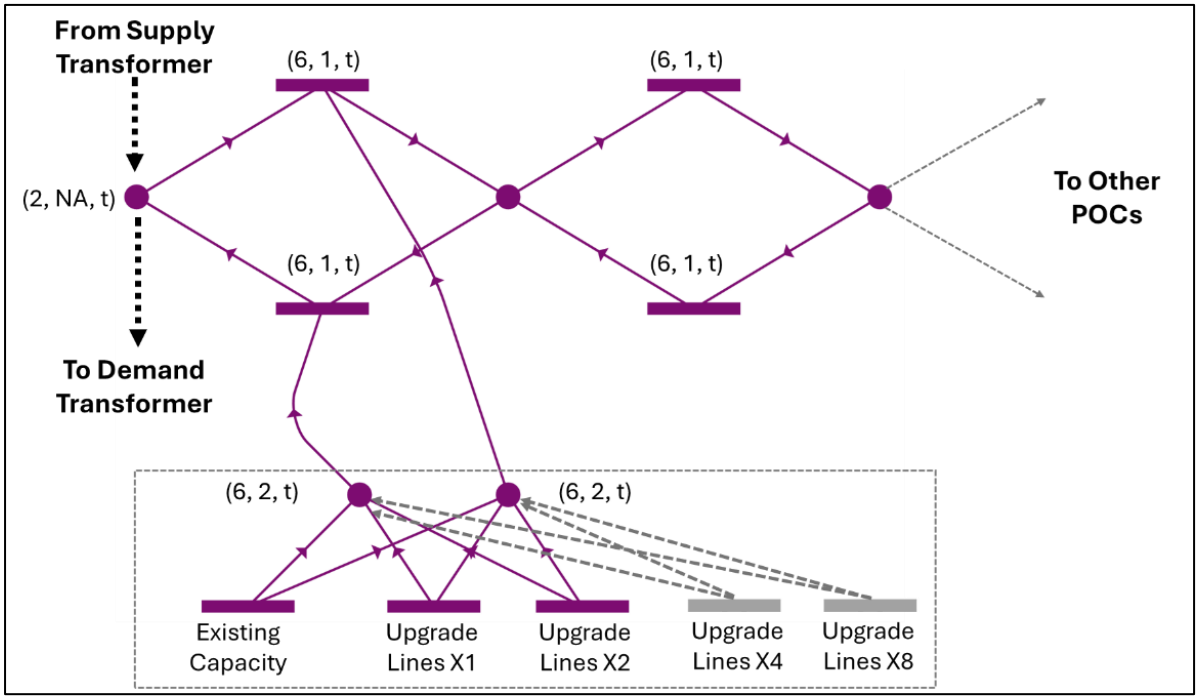


Figure F 9: Modular component for transmission upgrades when |T|=1.

As illustrated in Figure F 9, the model determines the number of identical grid reinforcements required by introducing a set of arbitrary potential upgrade lines. These lines are represented as operating units created by combining the identifiers of each POC pair to generate a unique label. Each upgrade option is then placed within an iterative loop indexed by n, where n denotes the number of parallel grid lines required to satisfy the load. The value of n may either be assumed or, more effectively, initialised using results from the simplified transmission grid modular structure. Once n is determined, the corresponding operating unit node is linked to a series of material nodes indexed by the same n, with these nodes varying across time periods. These material nodes then connect to their respective operating unit nodes; each is also assigned the specified n value. Through this arrangement, the model establishes the time-dependent energy flow pathways across the POCs, enabling the representation of transmission requirements as a function of the number of parallel grid lines needed to satisfy network demand.

Using this supervised benchmarking approach substantially reduces the number of variables and constraints in the full optimisation problem. In the New Zealand national case study, the

transmission dataset contains more than 600 POC to POC connections, even before accounting for multiple time periods. Benchmarking therefore provides an initial estimate of feasible capacities without imposing strict boundary constraints that could render the full model infeasible. This method enables a tractable and computationally efficient modelling process while preserving the ability to determine cost-optimal transmission upgrade pathways.

Transformers

The N security requirements for both demand and generation have to be established before the modular structure can be introduced. The transformer layer database contains the list of POCs, their electricity demand in MW, and the electricity supplied within each POC. Most specifications for the POC connections in the New Zealand case study are publicly available through Transpower and through the ERGO Consulting RETA reports. However, some privately owned POCs do not have publicly disclosed data. In these situations, transformer capacity is estimated based on the historical electricity demand recorded in the Electricity Authority database.

Figure F 10(a) illustrates the step-up transformer structure as represented in the P-graph model. The operating unit node with a middle identifier of 102 corresponds to the existing transformer capacity that enables electricity to be stepped up to the required transmission voltage. In parallel with this, a new operating unit node with a middle identifier of 57 represents the additional allowable electricity capacity made available through transformer upgrades. This upgraded capacity links to a designated operating unit layer, whose edges connect to a single material node with middle identifier 59. This node represents the total upgraded transformer capacity incorporated into the system. The edge between the operating unit and the capacity node includes a multiplier that accounts for the specified time resolution and undergoes the necessary conversion from MWh to GWh.

A similar structure is used to represent upgrades for step-down transformers (Figure F 10). These transformers are positioned between the POC transmission-level node and the distribution-level node. The step-down transformer operating unit shares the same middle identifier as the step-up transformer but uses a different prefix to distinguish the two functions. The upgraded step-down capacity is modelled in parallel with the existing step-down transformer unit. When the existing transformer capacity is exceeded, the upgrade operating unit with identifier 57 becomes active, ensuring that at least one of the capacity-governing nodes enters the solution. The mechanism introduces the appropriate additional cost incurred into the system whenever a new transformer capacity is required.

Both step-up and step-down transformers share the same cost parameters imported from the cost module. These costs are based solely on the installed capacity of the transformers, which determines the amount of electricity they can process. Detailed engineering design considerations such as operating voltage, current ratings, and frequency, leakage inductance and more are not incorporated into the cost calculation, since approximate cost scaling across the transmission grid is predominantly driven by apparent power capacity.

To reduce the size of the optimisation problem, the number of transformer upgrade iterations can be estimated by assigning an integer value to each POC. The required number of additional transformers can be approximated by summing the potential electricity generation and demand connected to each POC and comparing it with the maximum electricity required under a full electrification scenario. These estimate which can be generated within the database, are exported to the preprocessing model as initial parameter values. If such estimates are not supplied, the modular structure for each transformer operating unit defaults to an initialised integer value for the sets of transformer sizes applied uniformly across all POCs.

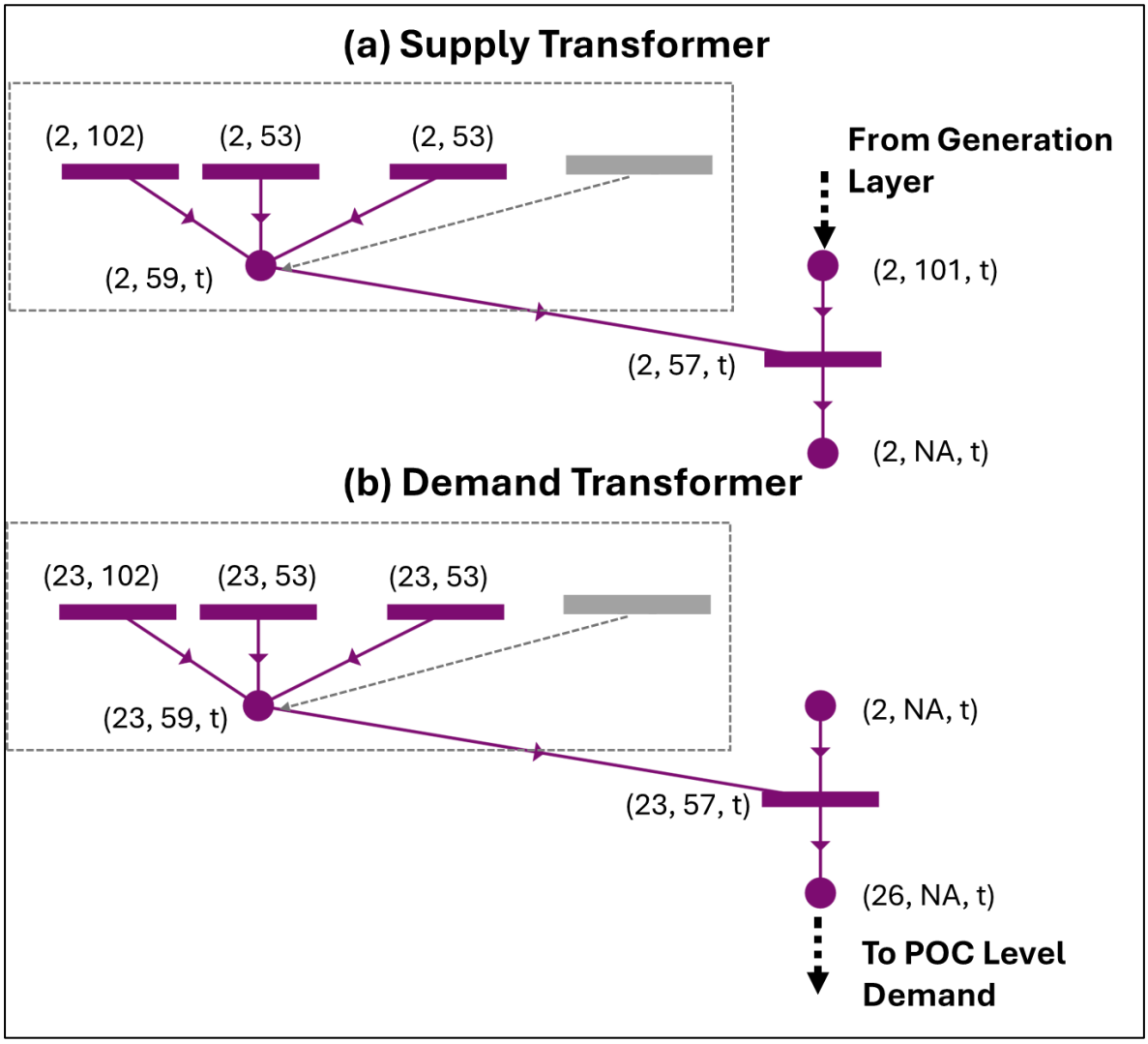


Figure F 10: Modular structure illustrating integer-based sizing of (a) supply and (b) demand transformer upgrades when |T|=1.

F2 Estimations on Residential, Commercial and BEV Demand

ChromeDriver automation for POC-level trading period datasets

From the 242 GXP points with a 30-minute interval for a 30-day month would end up being 10,454,400 entries, which would amount to a total of 381,585,600 entries for a single attribute, either electricity demand or wholesale electricity price. In a context of size, an Excel spreadsheet in 2010 has a limit of 1,048,576 rows and 16,384 columns. In order to combat these, the 242 GXP points can be simplified to 164 out of the 182 POCs in the processing step, where it is aggregated based on the service area for electricity distribution. Moreover, other challenges include the server storing the database having a limit for extracting the data, where the server connection timeout happens due to the download size exceeding the browser.

ChromeDriver is a standalone server that implements the WebDriver standard, which enables the user to automate its access to Chrome Browser from a Python script. ChromeDriver establishes the connection between the Python script and the Chrome Browser. Selenium is installed and imported as a Python library to be used as a WebDriver protocol so that the script would send instructions to open the server database and download the files with defined specific filters and attributes. For the electricity authority database, filters such as date range, region type (GXP nodes) and time scale are specified before triggering the data to be downloaded in a default CSV format. Once the download is completed, the Python script is set to automatically shift the date range forward by seven days. The cycle repeats itself by downloading the file sequentially from the 1$^{st\ of}$ January 2022 till the 31st of December, totalling 3 years' worth of data for every trading period for each GXP point. The downloaded CSV-formatted files are named sequentially for future processing.

Selenium interacts with the browser by using any test script from the client library, such as Python, Java or any other programming language, by transmitting commands (Bures and Filipsky, 2016). The transmitted commands are converted into a World Wide Web Consortium (W3C) format which is sent through HTTP requests to the Chrome Browser driver. The architecture of the ChromeDriver automation is shown in Figure F 11. The browser driver then receives the command, imitating a manual user to interact with the website. Once all files are downloaded, all individual files are organised and compiled into a single spreadsheet, with the entries to be transformed into rows of time periods and columns to be the POC.

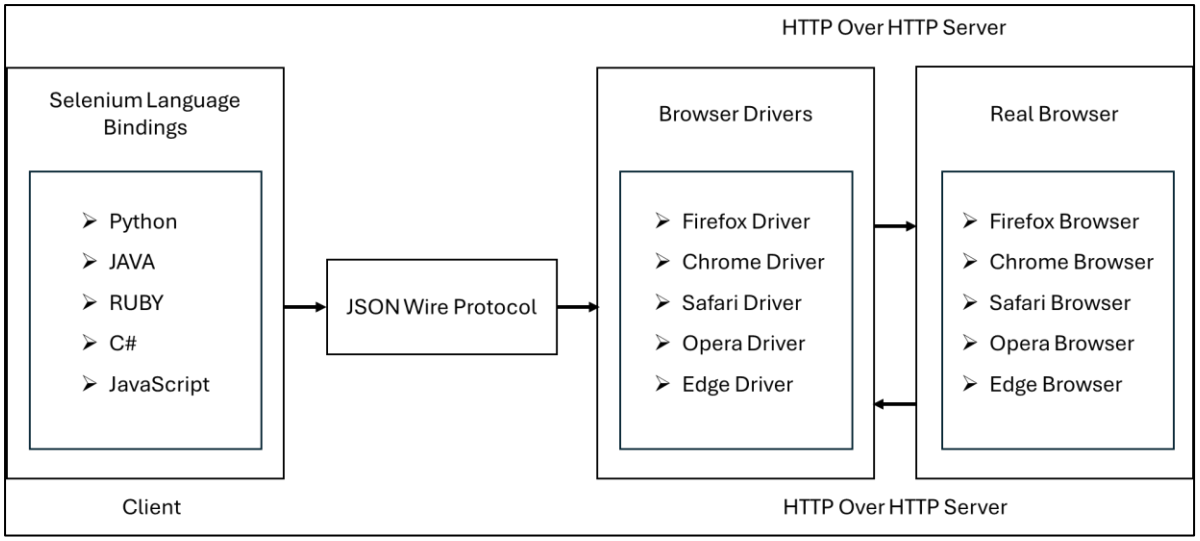


Figure F 11: Architecture of Selenium 3 WebDriver.

F3 Methods of Geocoding and Distance Calculation

F3.1 Geocoding of Factory and Generation Sites

As the process heat data originates from the Heat Plant Database and the RETA Database, one challenge is that the data for the Heat Plant Database does not have geometric coordinates, which would make energy flow from the resources difficult. Similarly, generation power plant datasets are obtained and compiled from the Electricity Authority, energy company websites, and the Open Infrastructure Map. Similarly, with the Heat Plant Database, there is only the name of the site and capacity without any geospatial element in the information. Therefore, geocoding tools involving Google Geocoding API and ArcPy are used where Google Geocoding returns the geographical coordinates and ArcPy measures the distance between each entry from two layers in the model.

Google's Geocoding API has been developed to be applied by inputting datasets containing names and adding new attributes such as coordinates to the table. Forward geocoding returns the latitude and longitude coordinates of the sites using the names or landmarks as the prompt. The Google Maps Platform was selected as it responds to real-time life updates, such as returning error messages when the location could not be found due to relocation or decommissioning. In order to minimise the error of incorrect matches due to similar street names used in other countries, the spatial search extent was constrained to the region so that all iterated entries will obtain their respective coordinates without any error.

The geocoding tool was developed in Python. A Google Cloud project is first created, then a unique API key is obtained by enabling the Geocoding API. Each request can be made by iterating through each entry for both generation and factory sites using a Python script which manages the client's API key and request handling protocol. The format returned by decimal degrees or the World Geodetic System 1984 (WGS84) which is a datum and geographic system, is then converted

to NZGD 2000 New Zealand Transverse Mercator (NZTM2000) format, which is a projected coordinate system. This measure is done by importing a geospatial-based analysis library package from Python called pyproj library to conduct this format change. The initial WGD84 which is WKID 4326 is converted to NZTM2000, WKID 2193.

F3.2 Spatial Allocation of Factory and Generation Sites to POC

Once the factory layer and generation layer contain geospatial attributes, these layers can be converted into point feature layers in SHP format to obtain the missing attributes for connected POC. For entries where the connected POC is not available in the source data, the model assigns the nearest POC as a proxy connection. This assumption is used to approximate the most plausible grid connection by minimising the required distribution or sub-transmission distance, and therefore the implied connection cost for electricity delivery.

After conversion to shapefiles, both layers are imported into ArcPy and the Near tool is applied to identify the closest POC for each factory site and generation site. The Near tool appends two key outputs to each input feature: (1) the straight line (Euclidean) distance from the input feature to its nearest POC, and (2) the identifier of that nearest POC. These attributes are then carried forward as the default POC connection for spatial allocation of demand and generation in the electricity system representation. As a result, the process heat and generation layers inherit both the assigned POC identifier and the corresponding POC distance attribute, which are then used as inputs for connection cost estimation and spatial allocation of demand and supply in the energy optimisation model.

F3.3 Distance estimation between Biomass and Factories or Generation Sites

Once each dataset layer has been identified and pre-processed with the required spatial attributes, it is necessary to define spatial relationships that act as proxies for transport connections. This is important because proximity between a biomass resource and a demand centre directly affects delivered fuel costs and therefore influences the optimisation outcome. This section outlines the ArcPy-based workflow used to estimate distance between biomass centres and factory or generation sites.

Distance Estimation using Constrained Road Networks

Biomass resource transport includes the application of road networks to the destination. Several studies apply Euclidean distance while estimating the distance between two coordinate points (Tatit, Adhinugraha and Taniar, 2024). However, as minimisation of the total system cost is the main objective in the energy system model, the results using Euclidean distance to correspond to delivery cost would be unreliable. This is because Euclidean distance would omit barriers such as protected areas, water bodies, mountain ranges and many more. Therefore, the distance calculation should be limited to a boundary that only includes service and public roads. The transportation layer represents topologically connected arterial routes which return the shortest distance path as raster data containing the cell's location. Essentially, the ArcPy tool does not natively support calculating overlapping distances between two points located on a polygon layer.

To address this limitation, a spatial barrier similar to Figure F 12 was constructed by overlaying the map of New Zealand with an erase feature. This step effectively creates the spatial extent within which distance calculations could be performed. This barrier serves as a constraint on the road network analysis, separating the North and South Islands, which ensures that the distance measurements between points are computed only within the valid geographic boundaries.



Figure F 12: Erased barrier polygons for road networks generated using ArcGIS Pro.

Distance Accumulation Tool

Distance Accumulation tool from ArcPy is used to return the raster details of travel distance from the source point within the constructed barrier. As the distance accumulation tool operates at the layer level rather than on individual entries, an iterative loop process has to be performed in Python to calculate the distance for each feature individually. Figure 6.23 shows the results of distance accumulation from a single source entry, whereas the colour gradient represents the distance with a cell size resolution of 400m^2. The final cell size is selected in ArcGIS through a trial-and-error approach.

A high cell size and barrier for roads in the pre-processing step will compromise the estimation of travel distance. This is because the distance accumulation algorithm may incorrectly identify

shortcut paths which do not accurately reflect the travel route and road condition within the road network. Conversely, a thin road network layer relative to the raster resolution can result in the loss of narrow road corridors during rasterisation. When the road width in the raster is smaller than the cell size, the road may not be represented as a continuous set of connected cells and may disappear entirely in some locations. Such an issue breaks network connectivity on the cost surface, so the distance accumulation algorithm cannot propagate through those sections, and the calculation may stop or return incomplete paths.