Decision-Centred Modelling for Industrial Process Heat Planning under Deep Uncertainty: Design Principles and an Edendale–Southland Proof of Concept

A thesis submitted to Auckland University of Technology

in fulfilment of the requirements of the degree of

Doctor of Philosophy

by

Seyed Ahmad Mahmoudi Lahijani

Supervisors: Dr. Mike Protheroe, Dr. Michael Gschwendtner

Department of Mechanical Engineering

School of Engineering, Computer and Mathematical Sciences

Auckland University of Technology (AUT) New Zealand

Kia whakatōmuri te haere whakamua

“I walk backwards into the future with my eyes fixed on my past”

— Whakataukī (Māori proverb)

Abstract

Industrial process heat is a pivotal yet under-served domain in energy transition analysis. Its decarbonisation is shaped by heterogeneous, site-specific equipment and operating constraints, long asset lifetimes, and tight dependencies on regional electricity headroom, renewable fuel availability, infrastructure delivery and evolving policy. In New Zealand these challenges are amplified by strong regional asymmetries and a data landscape that is improving in coverage but remains largely annualised and fragmented. Decision makers therefore face high-stakes, path-dependent technology choices, such as electrification versus biomass, whose consequences cross the boundary of any single model.

This thesis develops Decision-Centred Modelling (DCM), a modular site-to-region framework in which the analytical boundary is set by the consequences that can change the comparison of alternatives, not by the physical extent of the system being modelled. DCM rests on four commitments: a decision-first boundary principle; a thin-waist architecture in which modules exchange a small set of versioned, acceptance-gated artefacts (notably DemandPack and SignalsPack), so that methods may differ inside modules while interfaces stay fixed; robustness evaluation across ensembles of plausible futures using regret, satisficing rate and win rate; and progressive, non-destructive refinement guided by the analysis's own diagnostics.

The framework is demonstrated through a proof of concept at Fonterra's Edendale dairy site in Southland. It is methodological rather than predictive. A synthetic hourly demand profile is met by proportional dispatch under an hourly first-law energy balance; the exported incremental electricity is screened against Grid Exit Point headroom and a reinforcement menu; and two 2035 pathways, electrification completion and biomass completion, are compared over 100 paired futures. Under biomass completion, the 97 MW reinforcement was sufficient in every future; under electrification, the least-cost option switched to 150 MW in 49 futures, a threshold-triggered consequence that a site-boundary analysis cannot represent. Biomass completion had the lower system cost in 93 of 100 futures, with the ranking driven mainly by relative fuel prices. The diagnostics also showed that the most influential inputs, biomass cost and the dispatch-driven electricity peak, were among the least evidenced, so the analysis identifies its own refinement priorities. The results are conditional on two technology pathways, first-law energy accounting, and synthetic demand and headroom.

Table of Contents

Chapter 1 Introduction 1

1.1 Background and motivation 1

1.2 Process heat decarbonisation as a decision problem under deep uncertainty 2

1.3 Limitations of prevailing practice 2

1.4 Aim, research questions, and thesis objectives 3

1.5 Thesis contribution and significance 4

1.6 Scope, boundaries, and interpretation guardrails 5

1.7 Thesis structure 5

Chapter 2 Process Heat in New Zealand: Sector Structure, Regional Asymmetries, and Policy Evolution 7

2.1 Introduction 7

2.2 Industrial process heat and the New Zealand context 9

2.2.1 Global relevance of industrial process heat 9

2.2.2 New Zealand's process heat landscape 11

2.2.3 Fuel mix and emissions trends 13

2.2.4 Boiler stock and infrastructure 16

2.2.5 Regional disparities and site-level challenges 18

2.3 Policy mechanisms and decarbonisation initiatives 22

2.3.1 Early attention: climate agreements (1990s–early 2000s) 23

2.3.2 Evidence base: heat-plant surveys and renewables focus (mid-2000s) 24

2.3.3 Rising awareness: critique and reorientation (2010s) 24

2.3.4 Policy moves: action plans and instruments (2019–2021) 25

2.3.5 Regional initiatives: RETA and beyond (2021–Present) 25

2.3.6 Policy trajectory 25
2.4 Synthesis: implications for transition and decision support 26

Chapter 3 Conceptual Background: Principles for Decision-Centred Modelling 30

3.1 Introduction 30
3.2 Models and modelling: foundational concepts 32
3.3 Philosophical background: truth, uncertainty, and decision 35
3.3.1 Classical foundations: order vs contingency 35
3.3.2 Early modern rationalism, inductivism, and probability 37
3.3.3 Nineteenth and twentieth century: formalisation and systems 38
3.3.4 Late 20th century: bounded rationality, wicked problems, and resilience 40
3.3.5 Deep uncertainty and the decision turn 41
3.3.6 Epistemic beliefs and modelling practice 43
3.4 Systems definition and systems thinking 44
3.5 Principles for next-generation modelling 46
3.6 Synthesis and implications for transition and decision support 48

Chapter 4 Modelling Paradigms for Industrial and Energy Systems: Strengths, Limits, and Decision-Scale Gaps 50

4.1 Introduction 50
4.2 Structural classification 52
4.2.1 Top-down approaches 52
4.2.2 Bottom-up approaches 54
4.2.3 Hybrid approaches 55
4.3 Methodological classification 58
4.3.1 Optimisation: energy as a decision problem 58

4.3.2 Simulation: energy as a dynamic process 60
4.3.3 Equilibrium: energy as a balancing system 62
4.4 Decision orientation classification 66
4.4.1 Descriptive orientation 67
4.4.2 Prescriptive orientation 68
4.4.3 Exploratory orientation 69
4.4.4 Backcasting orientation 70
4.4.5 Robust and adaptive orientation 72
4.5 Additional classification dimensions 74
4.5.1 Purpose and function 74
4.5.2 Spatial and sectoral resolution 74
4.5.3 Temporal resolution and horizon 75
4.5.4 Market behaviour 76
4.5.5 Data intensity, transparency, and openness 76
4.5.6 Emerging approaches 77

Chapter 5 Decision-Making under Deep Uncertainty: Foundations for a Decision-Centred Framework 81

5.1 Introduction 81
5.2 Types and sources of uncertainty 83
5.2.1 Types of uncertainty 83
5.2.2 Sources of uncertainty 84
5.3 Deterministic and optimisation-based approaches 86
5.4 Probabilistic and stochastic decision approaches 91
5.5 Decision-Making under Deep Uncertainty (DMDU) 95
5.5.1 Robust Decision Making (RDM) 95

5.5.2 Dynamic Adaptive Policy Pathways (DAPP) 98
5.5.3 Many-Objective Robust Decision Making (MORDM) 102
5.5.4 Complementary DMDU building blocks 105
5.6 Scenario discovery and backcasting 108
5.7 Multi-Criteria Decision Analysis (MCDA) 113
5.7.1 Value-of-Information (VoI) analysis 116
5.7.2 Participatory MCDA and Co-production frameworks 117
5.8 Socio-technical and behavioural extensions 119
5.8.1 System Dynamics (SD) 119
5.8.2 Agent-Based Modelling (ABM) 120
5.8.3 Stock–Flow-Consistent (SFC) 122
5.9 Implications for framework design 125

Chapter 6 Architecture of the Framework: Requirements, Interfaces, and Artefacts 128

6.1 Introduction 128
6.2 Architectural overview 129
6.3 System boundary and scope definition 132
6.3.1 Spatial scope and carriers 133
6.3.2 Temporal scaffold and decision epochs 134
6.3.3 Mapping decisions to boundaries and information exchange 137
6.4 Functional and analytical requirements 139
6.4.1 Deterministic feasibility, coordination, and scale (R1–R3) 140
6.4.2 Exploratory robustness and adaptive pathways (R4–R5) 142
6.4.3 Provenance and auditability (R6) 144
6.4.4 Evaluation readiness and comparability (R7) 145
6.5 Decision–Input–Output–Constraint (DIOC) Map 147

6.6 Data preparation and packaging 149
6.6.1 Decision-centred schema (top-down) 150
6.6.2 Source acquisition and normalisation (bottom-up) 152
6.6.3 Chronology and compression mapping 154
6.6.4 Artefact design and external linkages 156
6.6.5 Data flow and governance 160
6.7 Operational specification: capabilities, layers, and coordination 164
6.7.1 Analytical capability specification and scope 165
6.7.2 Functional layers and iterative coordination 166
6.7.3 Complexity management and compression policies 168
6.8 Governance, assurance, interoperability, and roadmap 171
6.8.1 Quality assurance, validation, and transparency 171
6.8.2 Interoperability and adapters 173
6.8.3 Risks, assumptions, and limitations register 177
6.9 Extensibility roadmap and transition to the proof-of-concept 180

Chapter 7 Proof-of-Concept Case Study: Edendale & Southland 182

7.1 Introduction 182
7.2 Case study justification 183
7.3 Case study context and data overview 185
7.3.1 Edendale site overview 185
7.3.2 Southland regional context 186
7.3.3 Data sources 188
7.4 PoC Design: decision framing, scope, and artefact discipline 189
7.4.1 Decision framing and epochs 189
7.4.2 System boundary and coupling rationale 191

7.4.3 Artefact model, run organisation, and acceptance gates 196
7.5 PoC site-layer implementation 199
7.5.1 DemandPack construction 200
7.5.2 Epoch-tagged utility portfolios and availability representation 203
7.5.3 Proportional dispatch and ledger accounting (cost, emissions, feasibility) 206
7.5.4 Exported coupling signal 209
7.5.5 Site-layer validation checks used as reporting gates 210
7.6 Exploratory layer: screening and robustness evaluation 212
7.6.1 Decision locus and decision objects 213
7.6.2 Thin-waist artefacts and regional signals (screening inputs) 214
7.6.3 Uncertainty design and paired-futures contract 215
7.6.4 Regional screening mechanics and robustness metrics 216
7.6.5 Site decision robustness overlay 221
7.7 Integration and reporting gates 223

Chapter 8 Results and Discussions 225

8.1 Introduction 225
8.2 Deterministic site-core results across epochs 227
8.3 Demand profile evolution (DemandPack outputs) 227
8.3.1 Proportional dispatch results across epochs 231
8.4 Thin-waist coupling artefacts 239
8.4.1 Exported incremental electricity trajectories 239
8.4.2 Headroom feasibility diagnostics 242
8.4.3 Grid interface screening 244
8.5 Grid robustness screen (RDM) results across futures 248
8.5.1 Strategy dominance and robustness under 2035_BB 249

8.5.2 Strategy dominance, tail risk, and switching behaviour under 2035_EB 251
8.5.3 Decision-relevant outcomes 256
8.6 Site decision robustness overlay 257
8.6.1 Overlay construction 258
8.6.2 Ensemble comparison 259
8.6.3 Flip futures and key drivers 262
8.7 Cross-cutting discussion 265
8.7.1 Thin-waist modularity and interpretability 265
8.7.2 Governed reproducibility 266
8.7.3 Decision relevance under uncertainty 267
8.7.4 Validity and PoC limitations 269
8.7.5 Synthesis: what the PoC now justifies as the next step 270

Chapter 9 Conclusions and Perspectives 273

9.1 Contributions of the thesis 274
9.2 Summary of key findings from the PoC 275
9.3 Implications for energy transition analysis and governance in New Zealand 276
9.4 Limitations and validity: what remains conditional and why this is constructive 277
9.5 Roadmap for progressive expansion: from PoC to an “alive” decision environment 278
9.6 Concluding remarks 279

References 281

Appendix 303

List of Figures

Figure 2.1 Process heat energy consumption by high level sectors (% of energy consumed 2019) [6] 11

Figure 2.2Energy consumption and GHG emissions of process heat by fuel type in 2019 [6] 14

Figure 2.3 Energy consumption and GHG emissions of process heat by fuel type in 2024 [7] ... 14

Figure 2.4 Sectoral Distribution of Process Heat Capacity (MW) 2021 [39] 19

Figure 2.5 Regional Fuel Use Distribution for Process Heat (MW) 2021 [39] 20

Figure 2.6 Timeline of Process Heat Reporting and Decarbonisation Policy in New Zealand 26

Figure 3.1 Representative layers in model development and application [123] 34

Figure 5.1 Role of the deterministic core within the decision framework 91

Figure 5.2 Probabilistic and stochastic inner loop (with L-shaped iteration) 94

Figure 5.3 Conceptual DMDU workflow for process heat decarbonisation 108

Figure 5.4 Conceptual pipeline for scenario discovery and backcasting 112

Figure 5.5 MCDA decision layer within the process heat framework 118

Figure 5.6 Socio-technical modules (SD, ABM, SFC) interfacing with the decision framework 124

Figure 5.7 Integrated, modular decision framework for process heat decarbonisation 126

Figure 6.1 Multi-scale architecture and thin-waist information flow 131

Figure 6.2 Nested spatial tiers with exemplar constraints and exchanged signals 134

Figure 6.3 Temporal scaffold and decision epochs 136

Figure 6.4 Data pipeline in a standard run 162

Figure 8.1 Daily heat-demand envelope in 2020 (MW) 228

Figure 8.2 Load duration curve in 2020 229

Figure 8.3 Monthly total heat demand in 2020 (GWh) 229

Figure 8.4 Mean diurnal profile by season in 2020 (MW) 230

Figure 8.5 Hourly heat-demand trace for a representative week in February 2020 (MW)..... 231
Figure 8.6 Unit-stack dispatch (2020)..... 233
Figure 8.7 Unit-stack dispatch (2025)..... 234
Figure 8.8 Unit-stack dispatch (2028)..... 234
Figure 8.9 Unit-stack dispatch (2035_EB)..... 235
Figure 8.10 Unit-stack dispatch (2035_BB)..... 235
Figure 8.11 Unit utilisation duration curve (2035_EB)..... 236
Figure 8.12 Unit utilisation duration curve (2035_BB)..... 237
Figure 8.13 Units online proportion (2035_EB)..... 238
Figure 8.14 Units online proportion (2035_BB)..... 238
Figure 8.15 Duration curves of exported incremental electricity across epochs..... 241
Figure 8.16 Summary comparison of exported incremental electricity metrics across epochs242
Figure 8.17 Edendale GXP headroom Ht duration curve (reference signal used for screening; 2035 headroom shown for illustration)..... 243
Figure 8.18 Representative-week time series of headroom Ht illustrating step-like intra-week structure (reference signal; 2035 shown)..... 244
Figure 8.19 Representative-week overlay of incremental electricity ΔEt and headroom Ht (2025)..... 245
Figure 8.20 Representative-week overlay of ΔEt and Ht (2028)..... 246
Figure 8.21 Representative-week overlay of ΔEt and Ht at the 2035 fork (EB vs BB): pathway separation under a common reference headroom signal..... 246
Figure 8.22 Auto-selected grid upgrade choice vs headroom across 100 futures (2035_BB)... 249
Figure 8.23 Annual shed by strategy across futures (2035_BB)..... 250
Figure 8.24 Auto-selected grid upgrade choice vs headroom across 100 futures (2035_EB)... 252
Figure 8.25 Annual shed by strategy across futures (2035_EB)..... 253

Figure 8.26 Cost–shed trade-off scatter by strategy (2035_EB); cost axis in NZD ($\times 10^{10}$); shed on a logarithmic axis 254

Figure 8.27 Mean annual shed under AUTO as a function of headroom and incremental-load multipliers (2035_EB) 255

Figure 8.28 Paired-futures system cost distribution (2035_EB vs 2035_BB); values in NZD ($\times 10^{8}$); boxes show the interquartile range, whiskers the full range 260

Figure 8.29 Regret CDF under the site decision overlay (2035_EB vs 2035_BB); regret in NZD ($\times 10^{7}$)..... 261

Figure 8.30 Winner map in price space (Pelecvs Pbio, coloured by preferred pathway) 264

List of Tables

Table 2-1 New Zealand Boiler database summary [7] 16

Table 4-1 Classification lenses in energy system modelling 51

Table 4-2 Comparison of structural paradigms in energy system modelling 57

Table 4-3 Comparison of methodological paradigms in energy system modelling 65

Table 4-4 Comparison of decision orientations approaches in energy system modelling 73

Table 4-5 Summary of extended classification dimensions 80

Table 5-1 Uncertainty Types and Sources in Process Heat Decarbonisation Modelling 86

Table 5-2 Complementary DMDU Building Blocks for Process Heat Decarbonisation 107

Table 6-1 Requirements, rationale, and indicative implementations 140

Table 6-2 Canonical indices and identifiers 151

Table 6-3 Normalisation conventions 152

Table 6-4 Quality Assurance (QA) gates for conformant inputs 153

Table 6-5 Artefact overview (purpose, core content, producer/consumer, QA) 156

Table 6-6 DemandPack schema (required and optional fields) 157

Table 6-7 SignalsPack schema (required and optional fields) 158

Table 6-8 OutlookPack schema (required and optional fields) 158

Table 6-9 Reserved expansion hooks (illustrative) 159

Table 6-10 Table 6.10 OutlookPack schema (required and optional fields) 160

Table 6-11 Acceptance gates 163

Table 7-1 XLRM framing of the proof-of-concept decision 190

Table 7-2 Boundary register for the 2035 decision 193

Table 7-3 Minimum PoC artefact set and acceptance gates (compact register) 199

Table 7-4 Site-layer reporting gates and file contracts 211

Table 7-5 Artefact table (screening inputs)..... 214

Table 8-1 Artefact table (screening inputs)..... 226

Table 8-2 Dispatch ledger summary (annual totals) 232

Table 8-3 Annual summary of exported incremental electricity (ΔEt) across epochs 240

Table 8-4 Reference headroom screening diagnostics based on hourly exceedance deficit

$Xt = (\Delta Et - Ht) +$ 247

Table 8-5 Grid robustness summary by strategy across 100 futures (2035_BB)..... 251

Table 8-6 Grid robustness summary by strategy across 100 futures (2035_EB) 254

Table 8-7 Fixed pathway artefacts re-costed by the overlay 258

Table 8-8 Uncertainty multipliers used in the paired-futures ensemble..... 259

Table 8-9 Site decision robustness summary across 100 paired futures..... 262

Table 8-10 Flip-region diagnostic and uncertainty salience (futures where EB is cheaper vs BB is

cheaper; sorted by impact on $\Delta Csys$)..... 263

Table 8-11 Canonical artefacts underpinning Chapter 8 claims 267

Attestation of Authorship

I, Seyed Ahmad Mahmoudi Lahijani, hereby declare that this thesis is my own work and that it has not been submitted, in whole or in part, for any degree or diploma at any university or other institution. To the best of my knowledge and belief, this thesis contains no material previously published or written by another person except where due acknowledgement is made in the text.

I further declare that all substantive contributions from others (including supervisory guidance, institutional support, and external data provision) are appropriately acknowledged, and that any third-party materials included in this thesis are used in accordance with applicable permissions and citation requirements.

Generative artificial intelligence (AI) tools were used in a limited, assistive capacity. AI-assisted search tools were used to help identify candidate literature; every source cited in this thesis was retrieved, read and verified by the author, and references were managed in Mendeley. AI tools were also used for language editing of drafted text and for assistance with debugging code; all code outputs were checked by the author through the acceptance gates described in Chapter 7. No AI tool was used to generate data, results or analyses. The research design, framework, modelling choices, interpretation and conclusions are the author's own, and the author takes full responsibility for the content of this thesis.

Seyed Ahmad Mahmoudi Lahijani

Publications

The systematic literature review undertaken for this thesis resulted in a peer-reviewed journal publication that developed a structured process for selecting modelling tools for industrial process-heat decarbonisation. While the full method and detailed results of that study are not reproduced in this thesis (to avoid duplication and manage scope), its insights informed the framing of modelling requirements and the design choices developed throughout the literature review and subsequent framework chapters. The framework developed in this thesis has subsequently been described in a preprint, and a further manuscript is under review; both are listed below.

Lahijani, A. M., Protheroe, M. D., & Gschwendtner, M. (2025). A modelling tool selection for decarbonising industrial process heat systems. Renewable and Sustainable Energy Reviews, 210, 115149.

Submitted: 28 February 2024_Revised: 20 November 2024_Accepted: 24 November 2024

https://doi.org/10.1016/j.rser.2024.115149

Mahmoudi Lahijani, A. (2026). Decision-Centred Modelling (DCM): A governed modular architecture for infrastructure planning under deep uncertainty [Preprint]. SSRN. https://doi.org/10.2139/ssrn.7240878

Lahijani, A. M. (2026). Use-relative governance of model-derived artefacts across modular modelling boundaries under deep uncertainty. Manuscript submitted to *Environmental Modelling & Assessment*.

Acknowledgements

I would like to express my sincere gratitude to my supervisors for their guidance, encouragement, and constructive feedback throughout this research. Their support has been essential in shaping both the direction and quality of the thesis.

I would also like to acknowledge the Energy Efficiency and Conservation Authority (EECA) for their support in providing access to the Regional Heat Demand Database (RHDD), which materially informed the context and development of this research.

I am deeply grateful to my wife, Parastoo for her patience, understanding, and unwavering support during the many phases of this work. Her encouragement made it possible to sustain the long and demanding journey of doctoral study.

I also wish to thank my mom and dad for their continued love, support, and confidence in me. Their belief and encouragement have been a constant source of motivation.

Finally, I acknowledge the wider community of colleagues and collaborators whose discussions and support have contributed to this work.

.

Abbreviations and Acronyms

Acronym	Definition
ABM	Agent-based model
CCC	Climate Change Commission (New Zealand)
CIP	Clean-in-place
CO_2e	Carbon dioxide equivalent
DAPP	Dynamic Adaptive Policy Pathways
DCM	Decision-Centred Modelling
DMDU	Decision Making under Deep Uncertainty
EB	Electrode boiler
EECA	Energy Efficiency and Conservation Authority (New Zealand)
ETS	Emissions Trading Scheme (New Zealand)
GIDI	Government Investment in Decarbonising Industry (New Zealand)
GXP	Grid Exit Point
IAM	Integrated Assessment Model
IEA	International Energy Agency
IPCC	Intergovernmental Panel on Climate Change
KPI	Key performance indicator
LP	Linear programming
MILP	Mixed-integer linear programming
MBIE	Ministry of Business, Innovation and Employment (New Zealand)
MORDM	Multi-Objective Robust Decision Making
NPV	Net present value
NZ	New Zealand

OPF	Optimal power flow
PoC	Proof of concept
PRIM	Patient Rule Induction Method
RDM	Robust Decision Making
RETA	Regional Energy Transition Accelerator (EECA)
RHDD	Regional Heat Demand Database (EECA)
ToU	Time of use
VOLL	Value of lost load

Nomenclature

This nomenclature summarises the principal symbols, indices, and thin-waist contracts used in the thesis. Unless otherwise stated, time is indexed at hourly resolution in UTC (Δt = 1 h), and units are given in square brackets.

Indices and sets

- $t \in \mathcal{T}$— Hourly time index; $t = 1, \ldots, 8760 [h]$
- $y \in \mathcal{Y}$— Epoch/year index (e.g., 2020, 2025, 2028, 2035) $[yr]$
- $e \in \mathcal{E}$— End-use category index (e.g., process heat, site electricity)
- $u \in \mathcal{U}$— Utility/technology unit index (e.g., coal boiler, electrode boiler, biomass boiler)
- $r \in \mathcal{R}$— Region / GXP / network node index (as applicable)
- $s \in \mathcal{S}$— Scenario / pathway index (e.g., EB, BB variants)
- $\omega \in \Omega$— Future/uncertainty realisation (RDM future)

Time series, demand, and scaling (DemandPack)

- $s_{t,e}$— Normalised hourly shape for end-use e(dimensionless)
- α_e— Scaling factor mapping a normalised shape to absolute demand for end-use $e [\mathrm{MW}]$
- $D_{t,e}$— Hourly demand for end-use $e$at time $t [\mathrm{MW}]$
- E_e^{ann}— Annual energy for end-use $e [\mathrm{MWh/yr}]$
- Δt— Time step length (1 hour) $[\mathrm{h}]$

Demand mapping (shape → absolute):

- $D_{t,e} = \alpha_e \, s_{t,e}$

Annual energy consistency (used to determine α_e):

- $E_e^{ann} = \sum_{t \in \mathcal{T}} D_{t,e} \; \Delta t$

Site dispatch quantities

- Q_t^{dem}— Site heat demand at time $t [\mathrm{MW_{th}}]$
- P_t^{dem}— Site electricity demand at time $t [\mathrm{MW_e}]$

- $q_{t,u}$— Heat output from unit $u$at time $t[\mathrm{MW_{th}}]$
- $p_{t,u}$— Electricity input/output of unit $u$at time $t[\mathrm{MW_e}]$
- $A_{t,u}$— Availability of unit $u$at time t(0–1) (dimensionless)
- η_u— Efficiency / conversion factor for unit u(dimensionless)

Feasibility and shortfall accounting

- U_t— Unserved heat at time $t[\mathrm{MW_{th}}]$
- $E^{unserved}$— Total unserved energy $[\mathrm{MWh_{th}}]$
- $H^{unserved}$— Total unserved hours [h]

Unserved aggregation:

- $E^{unserved} = \sum_{t\in\mathcal{T}} U_t\ \Delta t$
- $H^{unserved} = \sum_{t\in\mathcal{T}} \mathbb{1}(U_t > 0)$

Grid interface signals (site → grid)

- ΔE_t— Incremental electricity export/import attributable to pathway change at time $t[\mathrm{MW_e}]$
- H_t— Headroom / available capacity at the GXP (or node) at time $t[\mathrm{MW_e}]$
- X_t— Headroom exceedance at time t, e.g. $X_t = \max(0, \Delta E_t - H_t)[\mathrm{MW_e}]$

Costs and economic parameters

- C^{fuel}— Fuel cost $[\mathrm{NZD/MWh}]$
- C^{elec}— Electricity tariff/cost $[\mathrm{NZD/MWh}]$
- C^{capex}— Capital cost $[\mathrm{NZD}]$
- C^{fix}— Fixed O&M cost $[\mathrm{NZD/yr}]$
- C^{var}— Variable O&M cost $[\mathrm{NZD/MWh}]$
- $VOLL$— Value of lost load / penalty for unserved energy $[\mathrm{NZD/MWh}]$
- r— Discount rate (dimensionless)
- CRF— Capital recovery factor (dimensionless)

Emissions (if reported)

- EF_f — Emission factor for fuel f $[\mathrm{tCO_2e/MWh}]$
- E^{CO2} — Total emissions $[\mathrm{tCO_2e}]$
- p^{CO2} — Carbon price $[\mathrm{NZD/tCO_2e}]$

Robustness / RDM notation

- ω — A "future" (uncertainty realisation) from the ensemble Ω
- $Z(s,\omega)$ — Performance metric (e.g., total cost) for pathway s under future ω
- $\rho(s,\omega)$ — Regret of pathway s in future ω (dimensionless or in metric units)
- $R(s)$ — Robustness score for pathway s (e.g., percentile regret, max regret)
- τ — Satisficing threshold (metric-dependent)

Regret (generic form):

- $\rho(s,\omega) = Z(s,\omega) - \min_{s' \in \mathcal{S}} Z(s',\omega)$

Chapter 1
Introduction

1.1 Background and motivation

The accelerating climate crisis has placed energy-system transformation at the centre of mitigation strategies. Within this landscape, industrial process heat, thermal energy used in manufacturing and processing, remains among the most difficult domains to decarbonise because of entrenched fossil-fuel dependence, heterogeneous site configurations, and long-lived assets that lock in decisions over decades [1]. Globally, process heat accounts for approximately one-quarter of final energy demand and remains disproportionately supplied by fossil fuels, making it a persistent source of industrial emissions [2]. Recent estimates indicate energy-related CO_2 emissions of approximately 37.8 Gt-CO_2 in 2024, highlighting the continued difficulty of achieving rapid emissions reductions in heavy industry despite expanding low-carbon deployment [3].

In New Zealand, the challenge is amplified by the tight coupling between site decisions and regional constraints. Process heat contributes a substantial share of national emissions, reported as around 28% of energy-related emissions and approximately 8.3 Mt-CO_2e in preliminary mid-2025 reporting [4]. Demand is concentrated in energy-intensive sectors, particularly dairy manufacturing, pulp and paper, and food processing, with dairy alone contributing around NZ$20 billion per year in exports (approximately 3–4% of GDP), making it both an economic backbone and a focal point for credible decarbonisation pathways [5]. Yet the sector remains structurally “hard-to-abate”, shaped by sunk boiler infrastructure, regionally uneven fuel and electricity constraints, and policy and investment risks that can neither be ignored nor reduced to a single deterministic forecast [6].

A key practical difficulty is that the available evidence base, while expanding, remains fragmented across institutions and scales. The country now has multiple datasets and assessments spanning emissions inventories, boiler stock surveys, and regional transition studies [7], but these are often annualised and weakly linked to the operational time domain that determines peak stress, adequacy risk, and infrastructure exposure [8]. Regional transition work (e.g., regional assessments) provides valuable constraint narratives, while national datasets provide broad coverage of site-level totals; however, neither typically provides an auditable, time-resolved mechanism for translating specific facility pathway choices into quantified regional interface consequences under uncertainty.

This thesis is motivated by the need for a modelling environment that can operate credibly under these constraints: it must (i) remain interpretable and auditable when granular data are sparse, (ii) propagate uncertainty in decision-relevant forms, and (iii) evolve progressively as improved data and richer physics become available, without breaking comparability or provenance.

1.2 Process heat decarbonisation as a decision problem under deep uncertainty

Decarbonising process heat involves long-lived assets, sequential investment, and irreversible commitments. Decisions are typically path-dependent and early choices constrain future options through infrastructure lock-in, contracting arrangements, and production integration. Moreover, key drivers, such as, electricity prices and tariffs, biomass availability and delivered cost, carbon pricing and policy durability, consenting and delivery risks, and future production levels, often exhibit deep uncertainty (describes situations where decision-makers don't know or can't agree on models, probabilities, or outcomes for the future [9]). In such settings, it is not credible to assign stable probabilities to a small set of futures and optimise against a single expected case. Decision support must instead characterise robustness, which shows how strategies perform across a wide range of plausible conditions, where and why they fail, and what forms of “insurance” or flexibility reduce vulnerability [10].

A further complication is cross-scale coupling. A site pathway that is attractive under site-only accounting may impose incremental electrical load profiles that create binding constraints at the regional interface. Conversely, regional headroom limitations, reinforcement costs, or delivery delays can overturn the affordability or credibility of site electrification strategies. Robust planning therefore requires an analytical structure that can (i) preserve site operational realism at a decision-relevant temporal resolution, (ii) propagate consequences to the regional boundary using explicit interfaces, and (iii) evaluate strategies under uncertainty without losing interpretability or provenance.

1.3 Limitations of prevailing practice

Existing modelling and assessment practices provide valuable but incomplete support for industrial process heat transitions. National energy system optimisation models are powerful for economy-wide consistency and long-run investment logic, but they often struggle to represent facility heterogeneity, operational timing, and site-specific constraints at scale. Regional transition assessments can provide grounded narratives about constraints and opportunities, but they may report regional layers in isolation from facility-level operational signals, limiting the ability to quantify how specific site pathways translate into regional stress and investment exposure. Meanwhile, the data landscape,

while improving in breadth, remains fragmented across scales and frequently aggregated to annual totals, constraining validation and reducing confidence in time-resolved conclusions.

These limitations are not merely technical; they are governance challenges. Multi-layer studies often suffer from weak provenance: results become difficult to reproduce, audit, or extend because the chain from assumptions to figures is not controlled through canonical outputs and acceptance criteria. In decision contexts where stakeholders must justify investments, infrastructure planning, and policy design, “reproducible robustness” is not a luxury, it is a requirement for legitimacy.

1.4 Aim, research questions, and thesis objectives

The overarching aim of this thesis is to develop and demonstrate Decision-Centred Modelling (DCM): a modular modelling environment that links site-level process heat transition choices to regional constraints under deep uncertainty, while remaining transparent, auditable, and progressively improvable.

In DCM, the decision, rather than any single model, is the object of study. The analytical boundary is therefore drawn around the decision: it is the smallest analytical scope within which the consequences capable of changing the comparison of the available alternatives are visible. This thesis terms this the decision-first boundary principle. Models are instruments in the service of the decision, selected and replaced as evidence and modelling maturity allow, and coupled through defined, versioned interfaces.

This aim is pursued through four objectives:

1. **Conceptual development:** synthesise insights from energy modelling, decision science, and deep uncertainty traditions into design principles for next-generation models.
2. **Framework design:** design a modular, site-to-region architecture with explicit interfaces and canonical artefacts that stabilise information exchange across scales.
3. **Uncertainty-aware analysis:** embed exploratory ensemble methods and robustness diagnostics suitable for deep uncertainty, prioritising vulnerability characterisation and least-regret reasoning.
4. **Application and learning:** demonstrate the framework via a proof-of-concept and extract methodological and policy-relevant lessons that guide targeted calibration and future expansion.

The thesis is guided by the following research questions:

RQ1: What modelling requirements are most critical for evaluating industrial process heat transitions under real-world technical, spatial, and policy constraints?

RQ2: How can site-level technology and operational decisions be coupled to regional constraints in a manner that remains interpretable and auditable under limited data?

RQ3: How can deep uncertainty in key drivers (prices, supply, policy, infrastructure delivery) be propagated to decision-relevant outputs using robust decision-making concepts?

RQ4: What methodological lessons from a bounded proof-of-concept can inform a credible roadmap towards a progressively expanding, "alive" decision environment?

1.5 Thesis contribution and significance

This thesis contributes a structured alternative to monolithic, forecast-centric modelling for process heat, organised around the four design commitments of DCM (a decision-first boundary, thin-waist artefact coupling, robustness evaluation under deep uncertainty, and progressive refinement) and their demonstration. The principal contributions are:

- A decision-centred conceptual framing: treating the decision, not the model, as the object of study, so that the analytical boundary follows the consequences that can change the comparison of alternatives and positioning industrial process heat modelling as a robustness and learning problem under deep uncertainty, rather than a single-future optimisation exercise.
- A thin-waist, artefact-driven architecture: enabling site and regional modules to evolve independently while maintaining stable interfaces through canonical data products (e.g., DemandPack and SignalsPack); methods may differ within modules, while the interfaces between them are contract-governed.
- Governed reproducibility as a first-class design property: treating acceptance-gated artefact completeness, provenance, and traceability as prerequisites for reported results.
- A defensible proof-of-concept demonstration: showing that an end-to-end, uncertainty-aware modelling chain can produce interpretable, decision-relevant evidence even when operational data are sparse, provided boundaries are explicit and outputs are auditable.
- A roadmap logic for progressive refinement: demonstrating that improved data, richer physics, and multi-pass coupling can be introduced without disrupting the core logic, interfaces, or comparability of results.

The significance for New Zealand is practical; the framework is designed to complement existing national and regional studies by providing a transparent mechanism to translate facility pathways into time-resolved regional interface stress and to reveal which uncertainties truly govern preference reversals, thereby guiding targeted data collection and calibration efforts. More broadly, the thesis offers a replicable pattern for other data-constrained industrial decarbonisation contexts; begin with governed, auditable artefacts and progressively increase fidelity without sacrificing interpretability or stakeholder trust.

1.6 Scope, boundaries, and interpretation guardrails

A central methodological stance of this thesis is to make boundaries explicit, not implicit. The proof-of-concept does not claim a fully calibrated operational replica of an industrial site, nor a planning-grade regional electricity model. Instead, it is designed to demonstrate an end-to-end architecture with clear interpretation guardrails: time integrity and energy closure at the site layer; explicit definition of the site-to-region coupling signal; screening-level regional evaluation under uncertainty; and robustness comparisons under paired futures. These boundaries are not an evasion of realism; they are a deliberate maturity strategy that enables defensible insight today and targeted refinement tomorrow.

1.7 Thesis structure

The thesis is organised as follows:

- Chapter 2 establishes the New Zealand process heat context, including sector structure, regional asymmetries, policy evolution, and the implications for modelling requirements.
- Chapter 3 develops conceptual foundations for next-generation modelling, including transparency, legitimacy, multi-scale coupling, and deep uncertainty as design drivers.
- Chapter 4 reviews modelling paradigms for industrial and energy systems and identifies decision-scale gaps that motivate a modular, uncertainty-aware approach.
- Chapter 5 reviews decision-making under deep uncertainty (DMDU) frameworks (e.g., robust decision making and related traditions) and distils their relevance for industrial decarbonisation planning.
- Chapter 6 specifies the proposed architecture, requirements, interfaces, and canonical artefacts, defining the thin-waist contract for cross-scale coupling and governed reproducibility.

- Chapter 7 implements the proof-of-concept case study for Edendale and Southland, emphasising methodological demonstration, artefact governance, and one-pass coupling.
- Chapter 8 presents results and discussion, demonstrating interpretability of the deterministic site core, decision-relevant regional screening under uncertainty, and paired-futures pathway robustness comparisons.
- Chapter 9 concludes by synthesising contributions, clarifying what is established empirically by the PoC, and outlining a staged roadmap for progressive expansion into a richer decision laboratory.

Chapter 2
Process Heat in New Zealand: Sector Structure, Regional Asymmetries, and Policy Evolution

2.1 Introduction

Process heat, a critical yet challenging sector in the global energy transition, underscores the urgency of transforming energy systems to mitigate environmental impacts while sustaining economic development [11]. Thermal energy used in manufacturing, refining, and chemical processing [12] remains a major source of greenhouse gas (GHG) emissions due to its persistent dependence on fossil fuels [13]. Despite its scale and importance, process heat has historically attracted less policy and modelling attention than sectors such as electricity or transport [14], leaving important analytical gaps in decision-support and long-term planning.

In New Zealand, process heat is particularly critical because it is deeply embedded in export-oriented industries such as dairy, wood processing, and food manufacturing, which collectively underpin national economic competitiveness. Decarbonising this sector is complex: boilers and supporting infrastructure have long lifetimes; replacement is capital-intensive; regional disparities shape access to renewable fuels; and industry, government, and communities often have competing priorities [6]. Policy initiatives such as the Emissions Trading Scheme (ETS) [4], the Government Investment in Decarbonising Industry (GIDI) Fund [15], and the Energy Efficiency and Conservation Authority's Regional Energy Transition Accelerator (RETA) [16], have sought to accelerate the shift. Yet their effectiveness ultimately depends on the ability of decision-makers to navigate technological, economic, and institutional uncertainties that determine the feasibility and timing of industrial transitions [17].

Recent years have seen an expansion of sectoral and regional data sources, including emission inventories, technology assessments, and fuel supply forecasts. However, these datasets remain fragmented across scales, from site-level boiler records to national system projections. Furthermore, decisions in this domain are characterised by deep uncertainty, where future trajectories of carbon pricing, fuel availability, technology costs, or policy regimes cannot be assigned reliable probabilities [18]. Traditional modelling approaches, reliant on deterministic optimisation or narrowly defined probabilistic frameworks, struggle to capture this breadth of plausible futures. Their reliance on single

scenarios or central forecasts risks producing fragile insights in contexts where structural change, policy volatility, and disruptive innovation are the norm.

Addressing this requires a re-thinking of how industrial transitions are modelled. Instead of "*predict-then-optimise*", emerging approaches emphasise "*explore-then-choose*" [9]. Methods such as Robust Decision Making (RDM) and Dynamic Adaptive Policy Pathways (DAPP) explicitly account for uncertainty by stress-testing strategies against ensembles of plausible futures and sequencing policies adaptively in response to triggers and signposts [19]. These techniques are increasingly combined with deterministic optimisation to preserve analytical rigour while enabling exploration of resilience and adaptability. As later sections will discuss, such methodological developments are not merely technical innovations but reflect deeper philosophical debates about truth, uncertainty, and decision-making.

A further defining characteristic of next-generation approaches is their modular and extensible architecture. Rather than seeking a single comprehensive model, modular frameworks allow site-specific analyses to inform regional and national systems while remaining flexible enough to incorporate new datasets, computational techniques, or policy concerns over time. This modularity, explored further in the discussion of International Energy Agency (IEA) Project BlueSky model [20] and other international frameworks, underpins the design philosophy of this research. It enables integration across scales, from boiler dispatch at industrial plants to regional fuel balances and national decarbonisation strategies, while ensuring transparency, reproducibility, and openness for stakeholder engagement across industry and policy [21].

This chapter compiles and examines the New Zealand evidence base required for credible analysis of industrial process heat. It situates the sector within the national socio-technical system; characterises heat services by temperature bands; traces fuel-use and emissions trajectories; maps the boiler stock, age profile, and process integration; and interprets regional asymmetries in fuels, grid hosting capacity (GXP headroom), and biomass logistics as binding feasibility constraints. It then reviews the evolution and layering of policy instruments (carbon pricing, regulation, capital support, and regional coordination) and their interactions with sector structure, which shape incentives, technological and institutional inertial structures, and exposure to uncertainty. The chapter closes by distilling decision-relevant implications and setting out the analytical needs that motivate the modelling work that follows. Evidence is drawn from a targeted narrative review that triangulates official datasets and reports (EECA's RHDD and end-use series; MBIE statistics; Climate Change Commission advice; major TSO and sector studies) with peer-reviewed literature (Scopus and Google Scholar, 1990–2025; search strings including "industrial process heat", "New Zealand", "biomass", "electrification", "ETS", "GXP"), using inclusion criteria that prioritise New Zealand industrial heat, documented methods, and

quantitative results. Statistics are cross-checked across sources and uncertainties are noted; data provenance and a source catalogue are provided in the appendix.

2.2 Industrial process heat and the New Zealand context

Industrial process heat is a cornerstone of modern energy systems, underpinning production in sectors ranging from food processing to heavy manufacturing [12], [22]. Its decarbonisation is a global priority because it involves large, continuous heat demands, persistent reliance on fossil fuels, and long-lived, capital-intensive assets [22]. These characteristics make the sector both economically essential and technically difficult to transform, with investment cycles and process integration creating path dependency and technology lock-in.

Although the scale and structure of process heat vary by country, common barriers persist including entrenched fossil infrastructure, regional disparities in access to low-carbon fuels and enabling technologies, and uncertainty in investment conditions and policy trajectories [23]. These conditions mean industrial decarbonisation is not merely a technical substitution problem but a *wicked problem* in the sense of contested objectives, shifting framings, and no single "right" answer, hence the need for plural approaches to uncertainty (probabilistic risk where appropriate; exploratory, deep-uncertainty methods where it is not) and decision support that acknowledges bounded rationality and institutional constraints [24].

To examine New Zealand's transition opportunities and constraints, this section reviews the global relevance of industrial process heat, New Zealand's process heat landscape, fuel mix and emissions trends, boiler stock and infrastructure, and regional disparities and site-level challenges. It concludes with a synthesis that highlights the implications of these factors for the country's decarbonisation transition.

2.2.1 Global relevance of industrial process heat

Industrial process heat accounts for the majority of final energy use in manufacturing worldwide and a significant portion of global industrial CO_2 emissions [25]. It is indispensable for thermal operations such as drying, sterilisation, smelting, and chemical conversion across industries including food and beverages, pulp and paper, metals, ceramics, and cement [26]. This universality makes process heat both a backbone of industrial production and one of the most difficult areas to decarbonise.

Globally, fossil fuels, particularly coal and natural gas, remain dominant because of their energy density, compatibility with existing infrastructure, and cost advantages. Low-carbon alternatives such as electrification, biomass, solar thermal, and hydrogen are emerging, but their uptake remains constrained by technological maturity, cost, and highly context-specific supply chain barriers [12]. As a result, industrial heat systems continue to reflect a strong path dependence, where legacy infrastructure and long-lived capital assets shape the timing and direction of transition.

Several features make this transition particularly challenging:

- **Asset longevity:** Many industrial heating systems operate for decades, locking in emissions and reducing the frequency of decision points for replacement.
- **Bespoke system design:** Heat systems are often engineered to meet process-specific requirements, making retrofits or substitutions technically and economically complex.
- **Uncertain investment environments:** Volatile fuel prices, fluctuating carbon markets, and evolving climate policies create hesitancy for long-term investment in low-carbon alternatives.

From a modelling perspective, these structural features highlight the risks of oversimplification. Process heat demand is frequently represented at coarse levels of aggregation in global or national energy models, reducing computational burden but obscuring critical distinctions in temperature requirements, technology feasibility, and regional resource availability[2]. Such simplifications can understate both the heterogeneity of industrial heat use and the exposure of transition strategies to long-term risk. Recent initiatives, including the IEA's 2024 Net Zero Roadmap and update and revised Nationally Determined Contributions (NDCs), stress the need for higher sectoral, technological, and spatial resolution in industrial modelling [27].

The global experience therefore offers two lessons. First, process heat cannot be treated as a uniform or easily substitutable demand category, and credible modelling requires segmentation by temperature, technology, and regional context. Second, because industrial transitions are subject to deep uncertainty, spanning technology costs, policy trajectories, and resource availability, robust decision frameworks that move beyond single-scenario optimisation are essential. These insights provide a critical benchmark for New Zealand, where similar structural barriers are compounded by distinctive regional energy profiles and industrial concentrations. Addressing these challenges requires frameworks that integrate detailed site- and regional-level representations with uncertainty-aware decision support to ensure that long-lived investments remain viable under diverse futures.

2.2.2 New Zealand's process heat landscape

Industrial process heat in New Zealand is overwhelmingly concentrated in the manufacturing sector, which accounts for approximately 81% of total consumption, reflecting the country's strong reliance on food and primary processing industries. As shown in Figure 2.1, the remainder is distributed across the commercial (10%), public (6%), and agricultural (4%) sectors [7]. This profile underscores the disproportionate role of industry compared to other users, reinforcing why decarbonisation of process heat is a central challenge for meeting national climate goals.

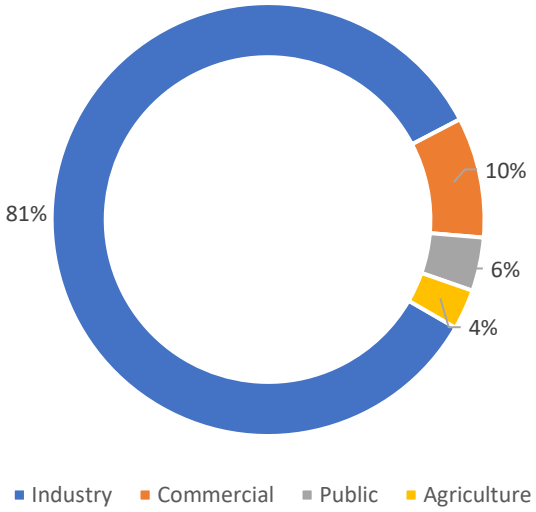


Figure 2.1 Process heat energy consumption by high level sectors (% of energy consumed 2019) [6]

A defining characteristic of New Zealand's process heat profile is its temperature distribution. Approximately two-thirds of total demand is for low- to medium-temperature applications (below 300 °C), typical of processes such as pasteurisation, sterilisation, and drying in dairying, meat processing, and pulp and paper production. The remaining demand, for high-temperature applications (above 300 °C), is concentrated in heavy industries such as metals, cement, and ceramics, where viable low-carbon substitutes remain limited [28].

From a decision-making perspective, this segmentation creates asymmetric challenges for transition pathways. Low- and medium-temperature services (e.g., hot water generation in food processing, pasteurisation in dairying, or drying in pulp and paper) are, in principle, compatible with a broad set of low-emissions alternatives [1]. Electrification through heat pumps and electric boilers, biomass combustion, and solar thermal systems are all technically feasible in these ranges [29]. However, uptake in practice is shaped by interrelated factors such as capital replacement cycles, site-specific

process integration requirements, and local resource availability (e.g., sustainable biomass supply, electricity grid hosting capacity). Even where alternatives are technically ready, firms often delay adoption until assets reach end-of-life or policy signals strengthen, creating transition inertia [30].

Temperature matters for more than technology screening, because it sets the quality of the energy a service requires. The first law of thermodynamics conserves the quantity of energy; the second law ranks its quality. The maximum useful work obtainable from heat Q supplied at absolute temperature T, relative to ambient temperature T_o, is its exergy, $Q(1 - T_o/T)$ [31]. Heat at around 180 °C therefore carries exergy equal to only about one-third of its energy content, whereas electricity is pure exergy. Combustion and resistive (electrode) heating both degrade high-quality inputs to lower-temperature heat, while heat pumps upgrade ambient or waste heat and deliver more heat than the electricity they consume, with a coefficient of performance bounded by the Carnot limit T_sink/(T_sink − T_source). Commercial high-temperature heat pumps now supply heat at 90 °C and above, and research units have reached about 160 °C [32] ; for European industry, around 78% of energy demand has been assessed as electrifiable with established technologies [33] . The physical principles and technologies of electrically driven process heating are set out in [34] . For modelling, the implication is that the temperature profile of a site's heat demand, not only its annual quantity, determines which electrification options are feasible and how much electricity they draw.

High-temperature services (e.g., cement clinker production, steelmaking, or ceramic firing) present more complex challenges. The range of viable substitutes is narrower, more costly, and technologically less mature [35]. Hydrogen combustion, electrified furnaces, and concentrated solar thermal are potential pathways, but their deployment depends on parallel infrastructure development and uncertain cost trajectories [29]. For instance, hydrogen uptake hinges not only on electrolyser cost declines but also on regional storage and distribution networks, while electrified furnaces would require major grid reinforcements. These uncertainties amplify investment risks, reinforcing a "wait-and-see" posture among firms.

In the context of robust decision-making, these sectoral and temperature-based distinctions carry several implications for modelling:

- The diversity of feasible options for low- and medium-temperature services makes scenario ensembles and sensitivity analysis essential to test how changes in costs, policy, or resource availability influence adoption outcomes.

- The persistence of long-lived industrial assets necessitates modelling of capital stock turnover and retirement inertia, whether stylised through aggregated representations or explicitly tracked in stock models.

- Where multiple technology options remain viable, multi-criteria evaluations (balancing cost, emissions, energy security, and social acceptance) are critical for identifying least-regret pathways.

These considerations highlight that New Zealand's industrial process heat cannot be credibly addressed through coarse or single-scenario optimisation alone. Instead, modelling frameworks must explicitly account for heterogeneity, path dependence, and uncertainty, in order to generate insights that are both analytically rigorous and decision relevant. Having established the sectoral structure and temperature distribution, the next step is to examine the fuel mix and emissions trends, which shape the scale and opportunities for low-carbon substitution.

2.2.3 Fuel mix and emissions trends

Fossil fuels remain the dominant source of process heat in New Zealand, though their relative shares and associated emissions have shifted in recent years. In 2019, natural gas supplied approximately 38% of process heat energy but accounted for more than half of sectoral emissions, reflecting its extensive use in medium- and high-temperature applications. Coal, though contributing only 11% of energy, produced around 26% of emissions due to its high carbon intensity (roughly 2.5 times that of gas per unit of energy). By contrast, bio-based fuels such as black liquor and wood chips provided 23% of energy supply while contributing less than 2% of emissions, highlighting their role as inherently low-carbon options where locally available [6]. These 2019 patterns are summarised in Figure 2.2.

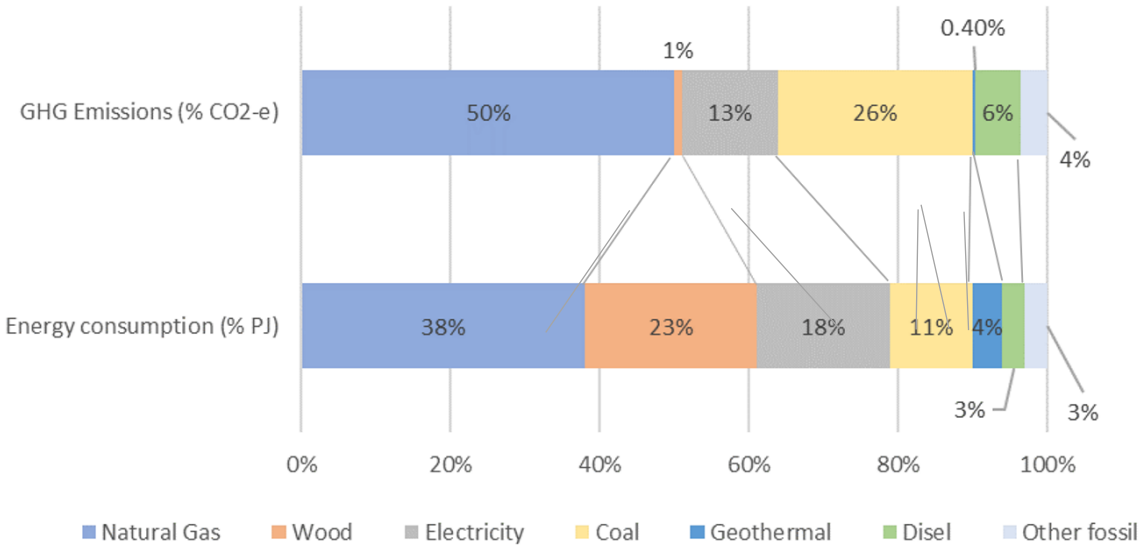


Figure 2.2Energy consumption and GHG emissions of process heat by fuel type in 2019 [6]

Recent data from EECA's *Regional Heat Demand Database* (RHDD) highlight an ongoing transition. As of 2024, fossil gas accounts for 52.9% of process heat demand, concentrated primarily in the North Island, and is responsible for 72% of sectoral emissions. Coal, used predominantly in the South Island, provides 12.4% of energy but 24.1% of emissions, reflecting its continued role in high-temperature applications. Bio-derived fuels now supply nearly 30% of energy demand but just 1.1% of emissions, largely concentrated in wood-processing industries where residues offer a natural feedstock [7]. These shifts are illustrated in Figure 2.3.

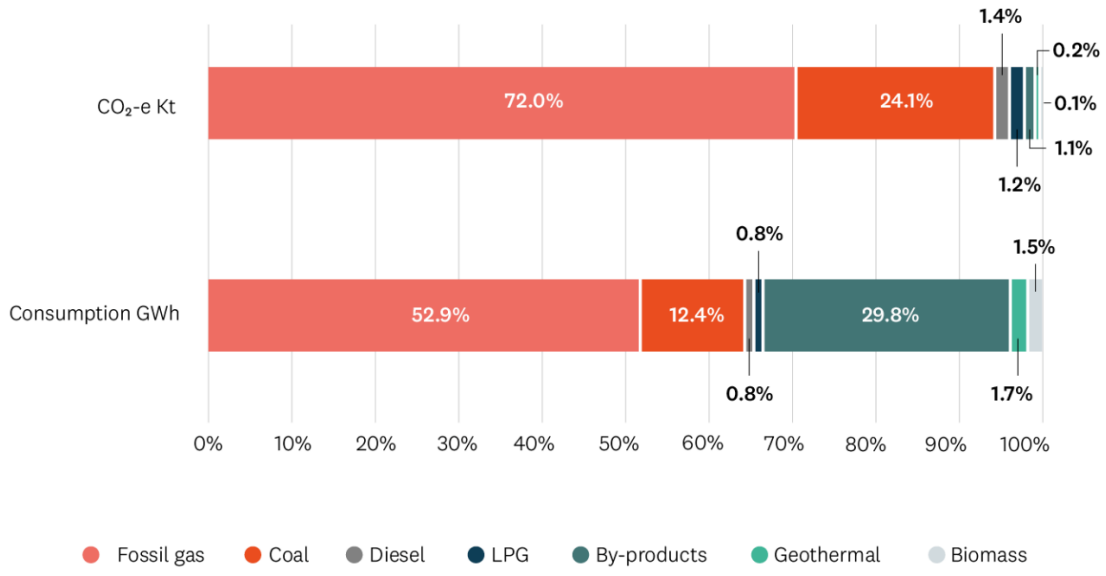


Figure 2.3 Energy consumption and GHG emissions of process heat by fuel type in 2024 [7]

These evolving patterns reflect the interplay of policy, infrastructure, and market dynamics. Gas remains dominant due to cost advantages, mature infrastructure (pipelines and boilers), and relative flexibility, despite its rising exposure to ETS carbon pricing. Coal persists where viable substitutes are scarce, such as cement kilns or metallurgical processes, illustrating the difficulty of replacing high-temperature fuels with currently immature alternatives like hydrogen or electric furnaces [36]. Biomass uptake has expanded in forestry-rich regions (e.g., Southland, Bay of Plenty) where transport distances are short and feedstock availability is high [37]. Yet, its national scaling potential is limited by geographic concentration, competing land uses, and sustainability constraints.

Policy instruments are shaping this mix but have not yet fully overcome structural inertia. The NZ ETS, with carbon prices reaching around NZ$100/t in mid-2025 [38], improves the competitiveness of biomass and electrification but remains subject to political volatility. The Government Investment in Decarbonising Industry (GIDI) Fund [15] has co-financed site-level conversions, especially to biomass and electric boilers, but aggregate impacts are constrained by capital lifetimes and site-specific integration challenges. The result is a sector still characterised by path dependency, with legacy assets and entrenched infrastructures dictating the speed and sequencing of change.

From a robust decision-making perspective, these dynamics translate into layered analytical requirements:

- **Near-term feasibility tests:** e.g., can a Southland dairy site economically shift from coal to pellets, or can an Auckland processor electrify steam production? This requires transparent evaluation of site-specific costs, resources, and integration challenges.

- **Sensitivity to volatility:** fluctuating gas prices, variable biomass yields, or uncertain electricity tariffs necessitate probabilistic sensitivity analysis to identify options resilient to shocks.

- **Long-term stress-testing:** strategic questions such as the timing of coal phase-out, the resilience of gas under tightening policy, or biomass sustainability under land-use competition demand exploratory scenario ensembles and stochastic optimisation. These approaches reveal whether planned pathways hold under disruptive futures (e.g., rapid electrolyser cost declines enabling hydrogen, or biomass shortfalls constraining regional strategies).

Taken together, these considerations show why least-cost optimisation alone is insufficient, while it can identify attractive pathways under baseline assumptions, it struggles with volatility and structural uncertainty. A more robust approach, combining cost testing, sensitivity analysis, and stress-testing across diverse futures, is required to design least-regret strategies for process heat transition.

However, these dynamics are mediated by the boiler stock itself; the physical infrastructure delivering most of New Zealand's process heat. Boiler age, efficiency, fuel flexibility, and retrofit potential all constrain the pace of substitution while shaping opportunities for staged or modular transition pathways. Understanding this stock is therefore essential for any credible modelling framework of industrial decarbonisation.

2.2.4 Boiler stock and infrastructure

Boilers remain the backbone of New Zealand's process heat system, supplying around 68% of national demand. Although typically designed for 15–20 years of operation, many units remain in service for 30–40 years, with capital replacement often deferred due to economic pressures, policy ambiguity, or the absence of clear decarbonisation signals [39]. This entrenched reliance on ageing assets creates both risks and opportunities: risks in terms of efficiency losses and emissions lock-in, and opportunities for targeted interventions when retirement or major upgrades become unavoidable.

The August 2024 update of the Regional Heat Demand Database (RHDD) records approximately 1,210 boilers across 670 industrial sites, with only 9% powered by renewable sources such as electricity or biomass [7]. These renewable systems provide about 1,470 MW of installed capacity, growth largely driven by GIDI-funded conversions [4]. By contrast, the remaining 1,090 fossil-fuelled units deliver more than 3,400 MW [7], underlining the scale of the challenge. Table 2.1 summarises the current boiler inventory, with minor adjustments based on EECA's 2025 projections.

Table 2-1 New Zealand Boiler database summary [7]

	South Island	***North Island***	***Total***
Number of Systems Identified	*460*	*750*	*1210*
Discrete Sites	*310*	*360*	*670*
Discrete Organisations	*230*	*190*	*420*
Total Installed Capacity	***1,850 MW***	***3,100 MW***	***4,950 MW***
Renewable Systems	*75*	*45*	*120*
Renewable Capacity	*370 MW*	*1,100 MW*	*1,470 MW*
Non-Renewable Systems	*385*	*705*	*1,090*
Non-Renewable Capacity	***1,480 MW***	***2,000 MW***	***3,480 MW***
Indicative Annual CO_2 Emissions	*3,700 kT CO_2 p.a.*	*3,800 kT CO_2 p.a.*	*7,500 kT CO_2 p.a.*

This stock of process heat assets reveals several decision-critical features that complicate decarbonisation pathways:

Asset heterogeneity: New Zealand's boiler fleet is diverse in size, age, duty cycle, and fuel type [7]. Some sites rely on small, intermittently fired boilers, while others depend on large baseload units

designed for continuous production. These differences influence efficiency, retrofit potential, and operational flexibility. For example, older coal boilers in Southland often operate below 70% efficiency, creating both emissions challenges and efficiency-gain opportunities through upgrades or control improvements. By contrast, modern gas boilers in Auckland achieve higher efficiencies but still leave scope for incremental measures such as condensate return systems or improved insulation [40]. This heterogeneity means generic prescriptions (e.g., “switch to biomass” or “electrify”) are rarely effective without detailed site-level appraisal.

Technology feasibility and sequencing: Replacement or retrofit potential differs substantially across sites. Some facilities may fully substitute fossil fuels. For instance, converting dairy dryers to electric boilers or biomass combustion. Others can only pursue incremental measures such as co-firing or installing waste-heat recovery systems [41]. Sequencing is critical; incremental upgrades (e.g., insulation, steam trap maintenance, or heat exchanger optimisation) can deliver 10–15% efficiency gains and reduce fuel use in the near term while paving the way for future transformative shifts [42]. Decision frameworks must therefore evaluate both incremental and transformational pathways and their cumulative effects over time.

Regional energy context: Transition options are shaped by regional disparities. Coal remains dominant in the South Island, where supply chains and boiler designs are entrenched, while natural gas infrastructure underpins many North Island clusters [16]. Regional conditions constrain options; firms in Waikato or Auckland are sensitive to gas and electricity tariff trajectories at their Grid Exit Points (GXPs) [43], whereas South Island firms face biomass logistics challenges and transport bottlenecks for pellet or chip delivery [44]. Regional grid capacity also limits electrification potential; some dairy processors cannot electrify boilers even with policy support due to GXP bottlenecks.

Uncertain timelines: Long asset lifespans and deferred replacement cycles compound uncertainty. Firms often delay retirement or investment due to unclear policy signals on coal bans, gas phase-out timelines, or carbon pricing [6]. Such ambiguity creates path dependency and lock-in risks: delaying retirement might allow adoption of more mature technologies in the future but also risks higher cumulative emissions and stranded assets if carbon prices rise sharply.

From a robust decision-making perspective, these dynamics suggest four analytical requirements:

- Constraint-aware evaluation at the boiler/site level, explicitly representing efficiency opportunities, technical feasibility, and duty cycles.

- Modelling of retirement inertia, recognising the importance of timing and sequencing in shaping emissions and costs.

- Mapping regional constraints, including grid hosting capacity and biomass logistics, to ground scenarios in realistic infrastructure conditions.

- Testing staged pathways under uncertainty, where incremental upgrades, partial substitutions, and eventual full replacements are compared under policy, price, and technology shocks.

Overall, New Zealand's boiler fleet defines both the technological backbone and the inertia of the process heat sector. Its heterogeneity, geographic distribution, and uncertain replacement timelines mean that credible transition strategies cannot rely solely on national averages or deterministic pathways. Instead, site- and region-sensitive approaches that stress-test multiple strategies under uncertainty are required to chart robust and least-regret decarbonisation pathways.

2.2.5 Regional disparities and site-level challenges

The spatial distribution of industrial sectors and energy resources in New Zealand creates pronounced regional disparities in process heat demand, fuel use, and emissions profiles. These disparities stem from legacy infrastructure, sectoral composition, and regional access to fuels, and must be explicitly incorporated into any credible decarbonisation pathway. Uniform national prescriptions risk overlooking the structural differences that define site- and region-specific transition opportunities.

The dairy sector dominates in both islands, with installed boiler capacity of approximately 630 MW in the South Island and 800 MW in the North Island [29]. While the North Island exhibits greater sectoral diversity, with food processing (260 MW) and wood processing (200 MW), the South Island's demand is concentrated in dairy and meat processing (190–200 MW). Smaller-scale users such as hospitals, schools, and municipal facilities contribute to distributed baseline loads, further complicating regional energy planning.

Figure 2.4 illustrates this sectoral distribution, showing both the dominance of dairy and the contrasting industrial profiles of the two islands. In the North Island, the broader mix of food and wood processing creates multiple transition entry points but also exposes firms to fuel-specific risks. In the South Island, reliance on dairy and meat intensifies vulnerability to disruptions in coal supply chains or biomass availability, with potential spillovers to export competitiveness.

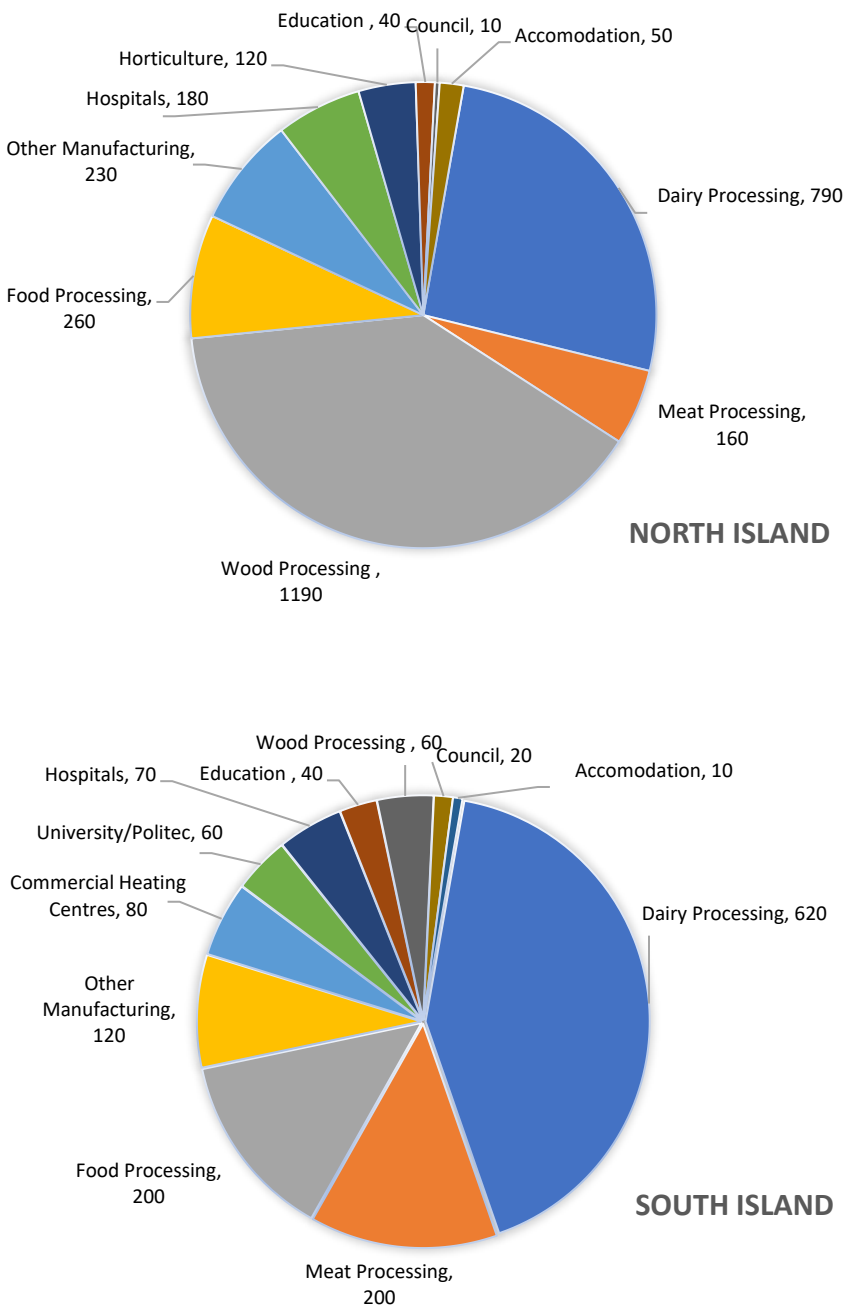


Figure 2.4 Sectoral Distribution of Process Heat Capacity (MW) 2021 [39]

Fuel use patterns diverge sharply between the islands (Figure 2.5). The South Island relies heavily on coal (about 1,100 MW), supplemented by diesel (270 MW) and biomass (330–350 MW), reflecting both historic investment in coal-fired boilers and the absence of reticulated natural gas. By contrast, the North Island is dominated by natural gas (1,750 MW), with growing use of biomass and electricity in wood processing and urban-based industries [29]. These dependencies embed structural risks, both physical (e.g., biomass supply chain fragility, grid bottlenecks) and policy-related (e.g., coal phase-out timelines, future of domestic gas supply) [39].

Figure 2.5 reinforces these asymmetries. Coal entrenchment in the South Island makes substitution costly and slow, while the North Island's gas reliance exposes firms to LNG price volatility and uncertain national phase-out signals. These divergent risk profiles highlight the limits of uniform national policy; strategies that appear cost-effective at aggregate level may prove fragile when applied to regional contexts.

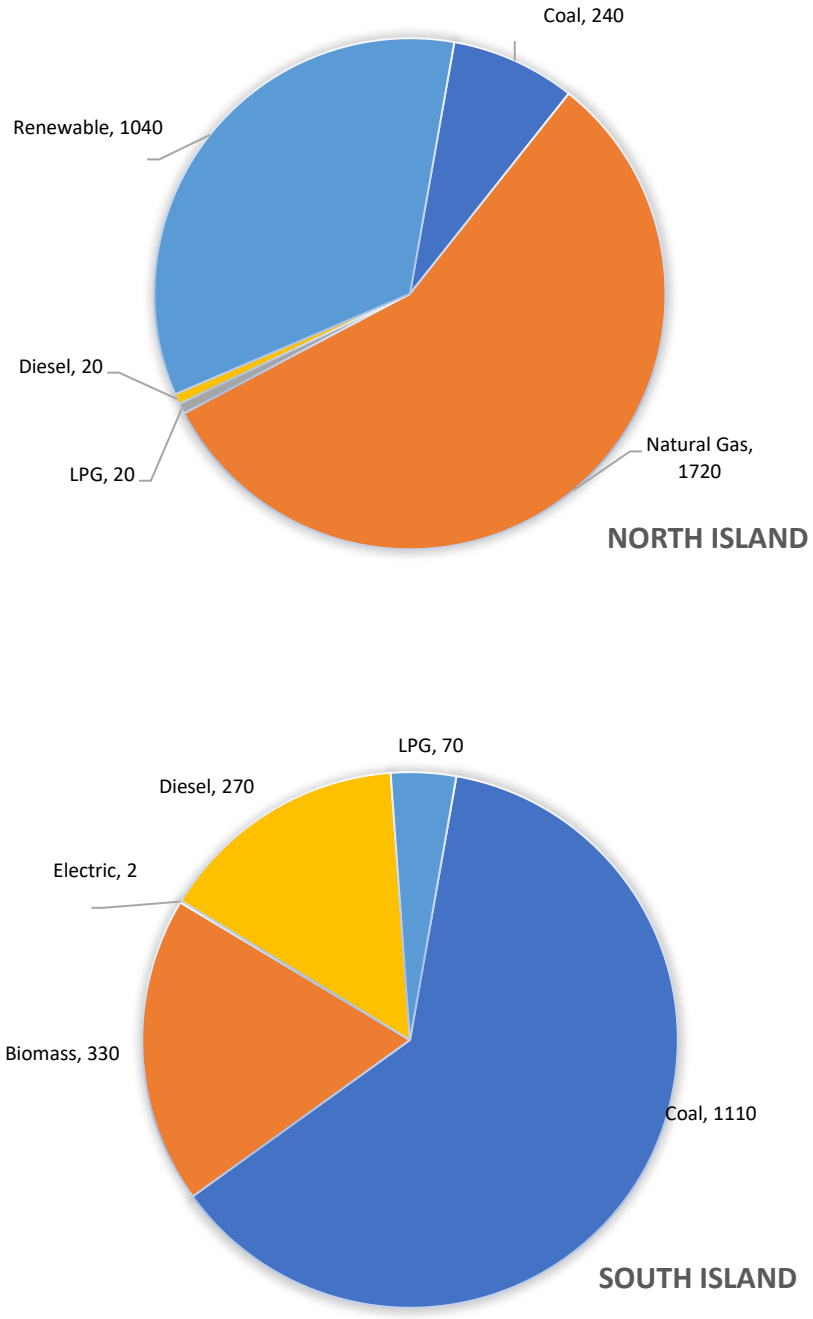


Figure 2.5 Regional Fuel Use Distribution for Process Heat (MW) 2021 [39]

Recent RHDD (2024) updates indicate modest change. Dairy sector capacity has increased slightly in both islands (to 630 MW in the South and 800 MW in the North), while wood processing capacity in the North has expanded by 5%, primarily due to improved use of bio-residues. Coal capacity in the South remains largely unchanged but is projected to decline by 5–10% by 2030, under the influence of the ETS and GIDI-driven conversions. Nationally, renewable capacity reached 1,500 MW in 2024, supported by a carbon price of NZ$100/t and increased policy momentum [4].

Despite this progress, several challenges persist:

- **Coal dependency in the South Island:** Long-standing supply chains, entrenched boiler infrastructure, and limited biomass alternatives make substitution slow and costly. Regional grid constraints further restrict the near-term scope for electrification. For example, dairy processors in Southland face both technical retrofit barriers and uncertainty over whether Grid Exit Points can support the additional load from large electric boilers [44].

- **Gas dependency in the North Island:** Gas remains critical for operational robustness and flexibility, particularly in industrial clusters around Auckland and Taranaki. However, this reliance exposes firms to LNG price volatility and uncertainty around the timing and scope of a national gas phase-out. A sudden withdrawal of gas supply, or a rapid price increase, could trigger stranded-asset risks if firms are locked into gas-reliant equipment [45].

- **Biomass supply fragility:** While biomass has expanded as a low-emissions option, its scaling remains vulnerable. In 2024, weather disruptions reduced supply availability by around 10 percent [46]. Beyond feedstock variability, constraints in storage, processing, and distribution infrastructure limit reliability at industrial scale[47]. For example, a food processor in Bay of Plenty may find biomass technically viable but commercially infeasible without parallel investments in local supply chains.

From a modelling perspective, these disparities highlight the necessity of regionally adaptive frameworks. Models must capture sector-specific demand structures, temperature requirements, and spatial fuel supply logistics (e.g., transport distances, infrastructure bottlenecks) [4]. They must also integrate regionally differentiated policy and market uncertainties. For example, Waikato and Auckland sites are highly sensitive to gas pricing and GXP tariffs [16], while South Island sites face biomass volatility and grid hosting limits for electrification [48]

Decision-analytic approaches provide useful lenses for these challenges. Scenario ensembles can stress-test strategies against biomass supply shocks or grid capacity constraints [49]. Sensitivity

analysis can explore industrial exposure to gas price volatility or ETS trajectories [4]. Adaptive policy pathways can help sequence incremental measures (e.g., efficiency upgrades, steam optimisation) before major technology shifts, reducing stranded-asset risks [50]. Multi-objective robust optimisation can balance cost, emissions, and reliability when evaluating site-to-region transitions [51].

In summary, New Zealand's regional disparities underscore the importance of moving beyond uniform national prescriptions. Coal entrenchment in the South Island, gas dependency in the North Island, and fragile biomass supply chains create regionally specific risks that cannot be addressed by one-size-fits-all solutions. Robust decarbonisation strategies require frameworks that explicitly integrate these structural variations, reflect long-lived asset constraints, and incorporate deep uncertainties into the decision space. While policy instruments such as the ETS, GIDI Fund, and RETA have provided important signals, their effectiveness depends on how sensitively they engage with these regional and sectoral realities. Therefore, the following section examines New Zealand's policy mechanisms as responses to these challenges, evaluating their contributions while identifying structural limitations that necessitate uncertainty-aware, next-generation modelling for facilitating robust decision-making.

2.3 Policy mechanisms and decarbonisation initiatives

New Zealand's efforts to decarbonise industrial process heat have been shaped by a diverse set of policy mechanisms introduced over the past two decades. These include market-based instruments such as the New Zealand Emissions Trading Scheme (NZ ETS) [4], direct regulatory interventions such as bans on new coal boilers [52], and targeted funding programmes like the Government Investment in Decarbonising Industry (GIDI) Fund [15]. More recently, regionally tailored initiatives such as the Regional Energy Transition Accelerator (RETA) [16], have sought to capture local conditions, addressing barriers specific to resource availability, grid capacity, and sectoral clustering.

Together, these mechanisms illustrate an evolving policy trajectory. Early measures emphasised broad efficiency improvements and voluntary uptake. Subsequent reforms shifted towards stronger interventions (such as, pricing, regulation, and subsidies) explicitly aimed at accelerating the phase-out of fossil-based heat [36]. Each instrument has been designed to overcome particular categories of barriers:

- The ETS targets economic barriers by altering relative fuel prices.
- Regulations address behavioural and lock-in barriers by prohibiting new high-emitting infrastructure.

- Funding schemes reduce upfront capital hurdles and share investment risk.
- Regional initiatives provide information and coordination where decentralised site-level data and decision-making would otherwise fragment transition pathways.

However, persistent challenges remain. Price volatility, long investment cycles, heterogeneous site conditions, and regional disparities complicate how these instruments influence actual decision-making. Policy effectiveness depends not only on the instruments themselves but also on how well they align with industrial realities and future uncertainties. These challenges exemplify process heat decarbonisation as a *wicked problem*.

For this reason, the assessment of policy mechanisms cannot be limited to static evaluation of costs and abatement potential. Instead, it requires decision-support approaches capable of accounting for uncertainty, path dependencies, and trade-offs. Methods such as Robust Decision Making (RDM), adaptive policy pathways, and multi-objective optimisation provide practical lenses for stress-testing policy impacts against fuel price volatility, biomass supply fragility, and grid bottlenecks [50]. These approaches can help identify where existing policies are effective, where they are insufficient, and what complementary measures are needed to support least-regret transition strategies.

This section therefore traces the historical development of New Zealand's process heat decarbonisation policy, from early strategies in the 1990s to the targeted interventions of the 2020s. It also assesses the structural limits of these instruments and highlights the methodological advances required to evaluate them under uncertainty.

2.3.1 Early attention: climate agreements (1990s–early 2000s)

Global climate agreements, particularly the 1997 Kyoto Protocol, played a formative role in placing industrial energy use, particularly process heat, on New Zealand's climate agenda [53]. The first National Energy Efficiency and Conservation Strategy (NEECS) in 2001 explicitly acknowledged the substantial share of industrial process heat in national energy use and outlined pathways to increase renewable heat sources such as biomass and geothermal [54]. It also anticipated the introduction of carbon pricing through future climate legislation, notably the Climate Change Response Act (2002), which implemented Kyoto commitments domestically [55]. While modest in scope, these measures established two critical precedents: recognition of process heat as a structural emissions source, and the embedding of carbon pricing into the long-term policy architecture.

2.3.2 Evidence base: heat-plant surveys and renewables focus (mid-2000s)

By the mid-2000s, process heat demand was estimated at approximately 90 PJ, with two-thirds supplied by fossil fuels [56]. Policy attention shifted towards developing a stronger evidence base. The Heat Plant Survey (2006) [57] and subsequent Heat Plant Database [58] catalogued hundreds of industrial boilers over 100 kW, mapping fuel use, age, and utilisation. These inventories represented a practical step towards addressing information barriers and enabled targeted policy design. Parallel strategies emphasised bioenergy and geothermal opportunities [59].

In 2008, the New Zealand Emissions Trading Scheme (NZ ETS) was legislated, extending to stationary energy in 2010. In theory, this placed a carbon price on fossil fuels used in boilers and should have incentivised fuel switching. In practice, however, the early ETS had little effect; free allocation of units and access to cheap international credits kept prices low [60], and by 2019, most large energy users reported that the ETS had not materially influenced investment in process heat [61], [62]. This revealed the weakness of relying on deterministic modelling assumptions that a single carbon price would directly drive behavioural change in a capital-intensive sector [63].

2.3.3 Rising awareness: critique and reorientation (2010s)

By the 2010s, process heat had become recognised as a "*hard-to-abate*" sector. The Royal Society (2016) reported that coal-fired heat alone contributed of around 8% of national emissions [64], while the Productivity Commission (2018) reported a 45% increase in process heat emissions between 1990 and 2016 [65]. This underscored the sector's lag relative to electricity.

Critiques of the ETS sharpened, where both modelling exercises and stakeholder feedback highlighted that low, volatile carbon prices were inadequate to drive major investment in low-emission heat technologies [63]. Policymakers began to introduce complementary measures. The 2017–2022 NEECS set the first explicit emissions-intensity reduction target for industry (1% per year) [66], and consultations between 2018–2019 marked a transition towards systemic policy engagement.

Sectoral reports reinforced this shift. Transpower's *Taking the Climate Heat out of Process Heat* (2019) assessed electrification potential [65]. Fonterra pledged to stop installing new coal boilers and phase out coal by 2037 [67]. The Interim Climate Change Committee (2019) recommended an immediate ban on new fossil-fuelled process heat systems [68], foreshadowing regulatory interventions. Academic and consultancy research (e.g., Motu [69], university studies [30], [70]) pointed to persistent barriers: high capital costs, limited biomass supply, and immature high-temperature alternatives.

2.3.4 Policy moves: action plans and instruments (2019–2021)

Momentum translated into action. MBIE's *Accelerating Renewable Energy and Energy Efficiency* (2019) [71], confirmed that over half of process heat demand was fossil-fuelled. Between 2020–2021, the Government launched the EECA-administered GIDI Fund, co-financing biomass and electrification projects [15]. A regulatory ban on new coal-fired boilers for sub-300 °C applications were signalled in 2020 and formalised in 2021 via a National Environmental Standard (NES) [72]. A phase-out of existing coal boilers for sub-300°C applications by 2037 was also announced.

In 2023, these measures were strengthened through updated National Environmental Standards (NES) on industrial GHG discharges [66], [73]. Climate Change Commission advice (2021) and the Emissions Reduction Plan (2022) prioritised process heat. Fonterra accelerated decommissioning of coal boilers, while NZ Steel entered a $140 million public-private partnership to install an electric arc furnace, cutting 800,000 t-CO_2 annually [74]. These developments illustrate how the groundwork of the 2010s translated into targeted action and investment in the 2020s, combining regulation and funding to drive change.

2.3.5 Regional initiatives: RETA and beyond (2021–Present)

In recent years, policy has moved toward regionally tailored delivery. EECA's Regional Energy Transition Accelerator (RETA) develops region-specific roadmaps that identify clusters of candidate conversions, estimate abatement potential and indicative costs, and map enabling infrastructure such as GXP reinforcements, feeder upgrades, and biomass logistics [16]. These reports combine industry engagement with screening-level techno-economic assessment to propose sequences of "no-regrets" and higher-impact projects, indicative timelines, and coordination needs with electricity distribution businesses and Transpower. The Auckland RETA (2021) examined electrification pathways and their network implications [75], while the Bay of Plenty RETA (2024) explored opportunities to anchor fuel switching with geothermal resources [37]. Collectively, these initiatives mark a shift from *why decarbonise* to *how delivery can occur in practice*, linking site actions to local grid capacity, fuel options, consenting pathways, and indicative investment windows.

2.3.6 Policy trajectory

Over the past two decades, New Zealand's policy response to industrial process heat decarbonisation has moved from broad efficiency initiatives to a more targeted mix of regulatory and financial instruments. Early inventories such as the Heat Plant Database (East Harbour Energy, 2008 [52]; CRL

Energy, 2011 [53]) provided the foundational evidence base, while subsequent interventions, including the Emissions Trading Scheme (ETS), the Government Investment in Decarbonising Industry (GIDI) fund, and the Regional Energy Transition Accelerator (RETA), sought to influence investment behaviour and to coordinate regional pathways. Figure 2.6 summarises this policy trajectory.

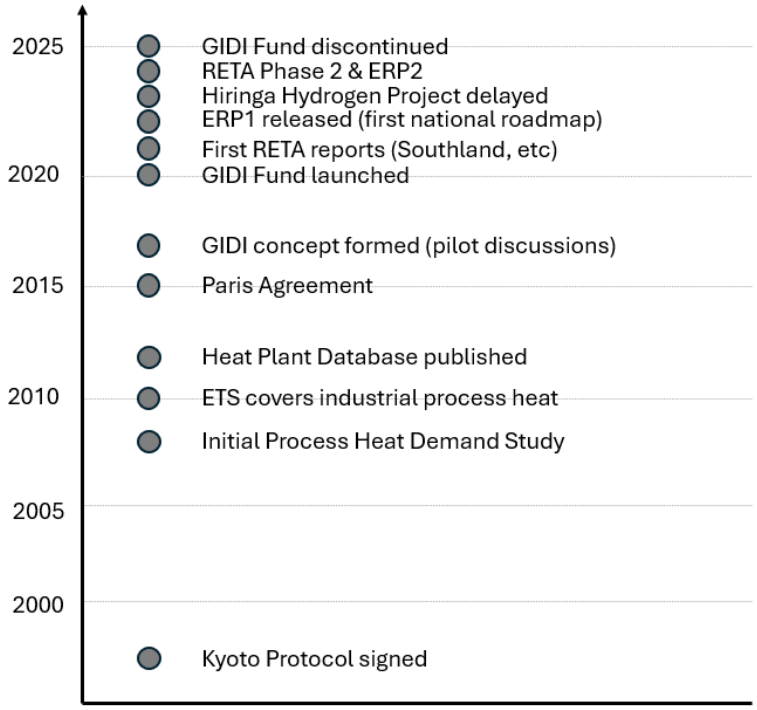


Figure 2.6 Timeline of Process Heat Reporting and Decarbonisation Policy in New Zealand

Taken together, this trajectory shows both progress and persistent shortcomings. Early reliance on price signals proved insufficient, prompting a gradual layering of regulation, subsidies, and regional planning. Yet systemic issues remain, including carbon-price volatility, uneven regional capacity, long investment cycles, and technology gaps. The next section (2.4) critically evaluates these structural and methodological limits, motivating robust, uncertainty-aware frameworks to strengthen New Zealand's process heat policy architecture.

2.4 Synthesis: implications for transition and decision support

The review of the global and New Zealand process heat landscape highlights several interconnected themes that underline the sector's complexity. Internationally, the persistence of fossil-dependent, high-temperature applications amid rising emissions limits the impact of incremental change. In New Zealand, this is amplified by spatial asymmetries (coal prevalence in parts of the South Island, gas

dependence in the North), heterogeneous boiler fleets and temperature-specific demands. These dynamics not only perpetuate emissions but also expose transitions to supply-chain risks and infrastructural bottlenecks, rendering decarbonisation far from a linear, like-for-like technical substitution.

Despite policy progress, systemic methodological weaknesses remain. Policy design has often reflected a "predict-and-plan" stance, relying on point forecasts of demand, prices and technology costs and then judging performance against those expectations [76]. In practice, the ETS has experienced price volatility and periods widely viewed as insufficient to drive fuel switching in process heat. At the same time, offset availability and sectoral exclusions have complicated incentives for direct industrial abatement [77], [78]. These dynamics expose the limits of static, deterministic price–behaviour assumptions under deep uncertainty.

A second theme is that plant-level efficiency does not automatically aggregate to system-level sustainability [79]. Industrial sites may meet international thermal benchmarks, yet upstream fuel risks (sourcing, logistics) and institutional lock-in (investment cycles, contracts, political economy) can undermine overall outcomes [80]. Studies report cases where high on-site efficiency coexists with environmental trade-offs or policy contestation, underscoring the need to look beyond narrow abatement metrics to regional resource constraints and governance realities [81], [82].

Third, programme design and delivery have not always aligned funding cadence, eligibility rules and project readiness [83]. Stop–start cycles and first-come allocation can privilege shovel-ready projects over system value, while uneven regional capacity affects sequencing and pace [84]. Regionally tailored efforts such as EECA's RETA are valuable for convening stakeholders and mapping options, but they face fragmented data and variable analytical practice, limiting integration with national planning. In particular, RETA's framing remains primarily diagnostic and facilitative rather than binding, with key limitations include:

(i) reliance on static hosting-capacity snapshots with limited treatment of time-coupled network constraints, coincident peaks and market operations;

(ii) coarse temporal resolution for industrial loads (annual energy/nameplate capacity) that obscures hourly diversity, thermal-storage value and demand response;

(iii) high-level treatment of biomass given uncertainties in sustainable supply, price formation and cross-sector logistics;

(iv) limited (and often qualitative) handling of uncertainty (ETS credibility and bands, technology learning, fuel prices, delivery risks) without robust decision metrics or adaptive triggers;

(v) data-transparency constraints (confidentiality, heterogeneous methods) that impede reproducibility and cross-regional comparability; and

(vi) partial coverage of delivery frictions (connection queues, consenting lead times, workforce and procurement capacity).

These gaps motivate a complementary modelling layer that couples site-level demand and utility dispatch with regional electricity and biomass at hourly (or representative-day) resolution and evaluates pathways under uncertainty, while retaining RETA's place-based focus.

Taken together, these patterns reveal structural and methodological limits in current decision-making approaches: an over-reliance on deterministic forecasts in contexts of deep uncertainty; assessment frames that privilege plant-level efficiency and near-term abatement while under-representing regional constraints, supply-chain risks and institutional factors; and fragmented interventions that are insufficiently integrated across sites, sectors and scales. These are not merely technical issues but reflect deeper epistemic assumptions and treating uncertainty as calculable risk rather than indeterminacy and viewing industrial transitions as linear substitution rather than "wicked" problems [85], characterised by ill-defined goals, stakeholder conflicts, and evolving interdependencies that resist definitive resolution. In the New Zealand context, "super-wicked" elements [86], including time-bound targets (e.g., net-zero by 2050) clash with short-term economic pressures, while actors contributing to emissions (e.g., dairy firms) are also tasked with solutions, further complicate matters. Conventional approaches, reliant on aggregated forecasts or deterministic optimisation, often overlook these facets, risking stranded assets or inequitable outcomes. This points to a paradigmatic shift from models assuming predictable pathways to frameworks that embrace deep uncertainty and adaptive planning [18]. Such frameworks must integrate systems thinking to address socio-technical interdependencies and modular designs to capture site-to-region dynamics [87]. International scholarship cautions that, under urgency, policy advances through path-dependent, incremental adjustments rather than clean structural breaks [88], [89]. In practice, small "paradigm nudges" [90], micro-transformations [91], or "soft transformative governance" [92], tend to erode lock-in overtime. New Zealand's process heat trajectory is consistent with this pattern, where progress has occurred, but largely through piecemeal, opportunistic interventions rather than a coherent, uncertainty-aware strategy.

For decision-making in New Zealand's industrial process heat decarbonisation, three key implications emerge. First, integration of heterogeneity and scale is essential. Analytical frameworks should link site-level realities (boiler configurations, operating envelopes, service levels) to regional resource and infrastructure constraints (GXP headroom and reinforcement timing, sustainable-biomass envelopes and logistics) and, in turn, to national objectives. Second, embedding uncertainty with decision-visible robustness metrics. Strategies should be stress-tested across ensembles of plausible futures (fuel/ETS bands, infrastructure timing, technology learning), with results reported in terms of reliability, satisficing shares and regret distributions, not solely single least-cost outcomes. Finally, balancing deterministic rigor with exploratory approaches. Optimisation remains critical for establishing engineering and economic feasibility, but it should be complemented by exploratory and participatory approaches that reveal vulnerabilities, trade-offs and adaptive sequencing through signposts and triggers.

These insights expose four interlocking gaps that shape the research need:

(1) A decision-scale methodological gap, as prevailing tools rarely combine hourly, process-aware site modelling with explicit regional feasibility checks in a single, auditable workflow.

(2) Uncertainty and robustness gap, with limited use of ensemble-based robustness methods and interpretable robustness metrics.

(3) Integration and traceability gap, reflecting weak "thin-waist" interfaces and incomplete provenance for reproducible coupling across modules; and

(4) New Zealand specific coupling gap, given the scarcity of open, reusable tools that reflect domestic realities, industrial heterogeneity, GXP-level constraints, biomass sustainability (and logistics), and evolving policy settings.

Addressing these gaps requires a framework that is modular by construction, coupled across scales (site → GXP/sub-region → region → national), uncertainty-aware via an ensemble runner with standard robustness reporting, and decision-visible, producing pathway artefacts (signposts, triggers) that support adaptive planning.

Chapter 3
Conceptual Background:
Principles for Decision-Centred Modelling

3.1 Introduction

Energy systems are complex socio-technical structures shaped by interactions among technology, markets, policy, and human behaviour [93]. Models are essential decision-support instruments in this setting. They are not neutral mirrors of reality but structured abstractions that simplify, formalise, and test assumptions to enable systematic analysis and decision-making. Properly designed, models can reveal underlying system dynamics, explore the consequences of alternative strategies, and illuminate trade-offs otherwise obscured in policy debate [94].

Energy modelling has historically been dominated by optimisation and simulation approaches, which offer structured ways to examine technology pathways, system dynamics, and policy options. At the global scale, integrated assessment models (IAMs) exemplify this tradition, underpinning Intergovernmental Panel on Climate Change (IPCC) assessments and IEA roadmaps by quantifying long-term emissions pathways and technology choices [95], [96]. While invaluable for framing international goals, these models have been criticised for their reliance on central forecasts and deterministic optimisation, which risk obscuring the deep uncertainties surrounding technology costs, behavioural adoption, and geopolitical change [97]. This critique catalysed approaches that test strategies across many plausible futures rather than predicting a single outcome; detailed frameworks are introduced in later chapters [9], [10].

The historical trajectory of energy modelling illustrates that paradigm shifts have often been triggered by major global political and economic events. Operations research emerged from wartime logistics in the 1940s and expanded in Cold War planning [96], [98]. Scenario planning gained prominence in the 1970s oil crises, most visibly through Shell's strategic foresight exercises [99], [100]. The 1990s saw the institutionalisation of the IPCC and the proliferation of IAMs [96]. More recently, resilience- and robustness-oriented approaches accelerated in response to the 2008 financial crash, the COVID-19 pandemic, and ongoing geopolitical disruptions [101], [102], [103]. These episodes underline that models are intellectual responses to crises as much as technical artefacts, reflecting the priorities and politics of their time [104].

The political–ideological backdrop has also shaped modelling practices. In the United States, modelling has often been embedded within a liberal-capitalist orientation, privileging optimisation and

forecasting as neutral instruments that legitimise market-led decisions [105], [106]. In Europe, by contrast, social-democratic traditions have fostered reflexivity, transparency, and open-source communities, even when the tools themselves remain optimisation-based [21], [107]. These ideological currents help explain the continued influence of IAMs and EIA/IEA forecasts in some contexts, and the greater traction of participatory and exploratory frameworks in others.

In recent years, the scope of modelling has widened through debates on *energy justice* and *the just transition*. Energy justice interrogates how the benefits and burdens of energy systems are distributed, the fairness of decision-making procedures, and the recognition of marginalised groups [108]. It extends evaluation beyond efficiency to distributive, procedural, and recognition-based justice, and increasingly to non-Western, intergenerational, and ecological perspectives. In parallel, the notion of a just transition, enshrined in the 2015 Paris Agreement, emphasises that decarbonisation must not reproduce inequality or leave workers and communities behind [109], [110]. For energy modelling, this implies moving beyond narrow cost-optimisation toward decision-support that foregrounds equity, legitimacy, and adaptability to evolving socio-political constraints.

In New Zealand, tools such as TIMES-NZ and initiatives like the Regional Energy Transition Accelerator (RETA) inform national and regional decarbonisation strategies [16], [111]. Yet diverse approaches across agencies and sectors have created fragmentation, with overlapping assumptions, limited transparency, and weak integration across scales [112]. This pluralism reflects the range of questions models are designed to answer, from national least-cost pathways to site-level technology feasibility, but it also underscores the challenge of producing consistent and actionable insights for industrial process heat, one of the country's most difficult decarbonisation domains.

The broader challenge extends beyond the energy field. Contemporary debates on large language models (LLMs) reveal similar issues of epistemic authority, bias, and institutional framing. Just as LLMs are shaped by training data and platform priorities [113], [114], energy models embed institutional assumptions (e.g., about costs, policies, and behaviours). Both cases illustrate the politics of knowledge production, where outputs reflect design choices and contexts as well as data.

This chapter therefore develops a conceptual synthesis linking foundational concepts of modelling, philosophical traditions of truth and uncertainty, and systems perspectives on interdependence and feedback. It begins by clarifying the role of models as structured abstractions and mediating instruments, then traces how differing epistemic commitments have shaped approaches to uncertainty and decision. A subsequent focus on system definition and systems thinking positions energy systems as socio-technical constructs, emphasising feedback, path dependence, and the need for site-to-region modularity. Building on these elements, the chapter distils key principles for next-

generation modelling. Taken together, these foundations motivate uncertainty-aware, multi-scale frameworks essential to industrial process heat decarbonisation and underpin the methodological framework developed later in the thesis.

3.2 Models and modelling: foundational concepts

A model is a simplified yet constructive representation of reality, designed to highlight essential dynamics while abstracting details deemed less relevant to the problem at hand. Stachowiak's classical framework defines models as entities that (i) map an original system, (ii) reduce complexity through abstraction, and (iii) serve pragmatic purposes [115]. In energy systems, models may take conceptual form (e.g., causal loop diagrams mapping feedbacks between fuel demand and emissions in dairy processing), mathematical form (e.g., optimisation problems for least-cost boiler dispatch), or computational form (e.g., simulation platforms for grid–biomass interactions). In each case, the goal is not replication but structured abstraction to support reasoning and decision-making [104], [116], [117]. In this thesis, a model is a transparent, problem-oriented abstraction that encodes explicit assumptions to generate decision-relevant outputs under clearly stated boundary conditions and data contracts.

A brief analogy with language and computation is instructive. Chomsky's theory of generative grammar conceptualises language as a formal system governed by simple principles that generate infinite complexity [118]. Similarly, energy models encode governing rules that generate a space of possible system behaviours. Recent debates on large language models (LLMs) caution that fluency does not guarantee truth; just as LLMs produce outputs shaped by training data and institutional contexts [113], [119], energy models generate results filtered through assumptions about technology readiness, carbon prices, or behavioural change. Both may appear internally consistent yet remain ungrounded in practical feasibility. The analogy is elaborated in this chapter and is used here to motivate transparency, and external validity checks rather than to equate modelling with language generation.

Morrison and Morgan (1999) characterise models as mediating instruments; autonomous entities that combine theory, data, and heuristics. They function as investigative tools, manipulated and stress-tested to generate insight rather than to reproduce truth [116]. Epstein (2008) similarly emphasises models as "generators of insight," guiding data collection, revealing uncertainties, and supporting structured exploration [117]. In energy transitions, this mediating role is crucial. Models provide laboratory-like testbeds to examine assumptions about demand growth, biomass supply, or grid expansion before costly infrastructure commitments. Their legitimacy depends not only on technical rigour but also on transparency of assumptions, provenance of inputs, and clarity of limitations.

Modelling is best understood as a process of framing and iteration, involving the definition of system boundaries, the formulation of assumptions, the choice of governing dynamics, and the validation of outputs against empirical evidence or expert judgement [104]. These decisions are never neutral. As Bock and Pfenninger-Lee (2025) observe, models embody epistemic positions about truth and certainty [120]. A model assuming a central planner allocating resources optimally reflects a positivist worldview, while exploratory scenario ensembles align with postpositivist and pragmatist traditions that emphasise pluralism, contingency, and decision-support. The politics of assumptions, whether about carbon prices, technology costs, or behavioural adoption, is inseparable from the knowledge claims that models generate.

Within the energy domain, three broad traditions of modelling can be distinguished:

- **Descriptive models:** map current states (e.g., energy balances, technology inventories) to establish baselines and accounting consistency.
- **Exploratory models**: test plausible futures via scenarios, sensitivities, or exploratory modelling (e.g., Shell-style narratives), prioritising breadth over point prediction.
- **Prescriptive models**: identify strategies deemed "optimal" under stated objectives and constraints through optimisation or multi-criteria evaluation (e.g., MARKAL/TIMES), subject to the assumptions embedded in problem formulation.

Each tradition reflects not only methodological choices but also deeper paradigms, which are structured ways of framing problems, selecting methods, and defining objectives [93]. In practice, many contemporary frameworks combine elements of these traditions (for example, embedding prescriptive optimisation within exploratory scenario ensembles), blurring boundaries while preserving distinct philosophical orientations. Paradigms evolve as societal concerns shift, from security of supply in the 1970s oil shocks, to global decarbonisation in the 1990s, to today's emphasis on resilience, justice, and robustness under uncertainty [121]. Importantly, paradigms accumulate rather than cleanly displace one another, reflecting the diversity of contexts in which energy models are deployed [122].

Figure 3.1 illustrates this layered structure of modelling [123]. On the data side, raw observations must be processed, harmonised, and formatted before becoming model inputs. On the structural side, governing dynamics are formalised mathematically and implemented in code. This layered view underscores how models mediate between raw data and abstract formulation: assumptions at any stage (temporal resolution, boundary definition, objective choice, or solver settings), influence not only technical plausibility but also the perceived legitimacy of results [124].

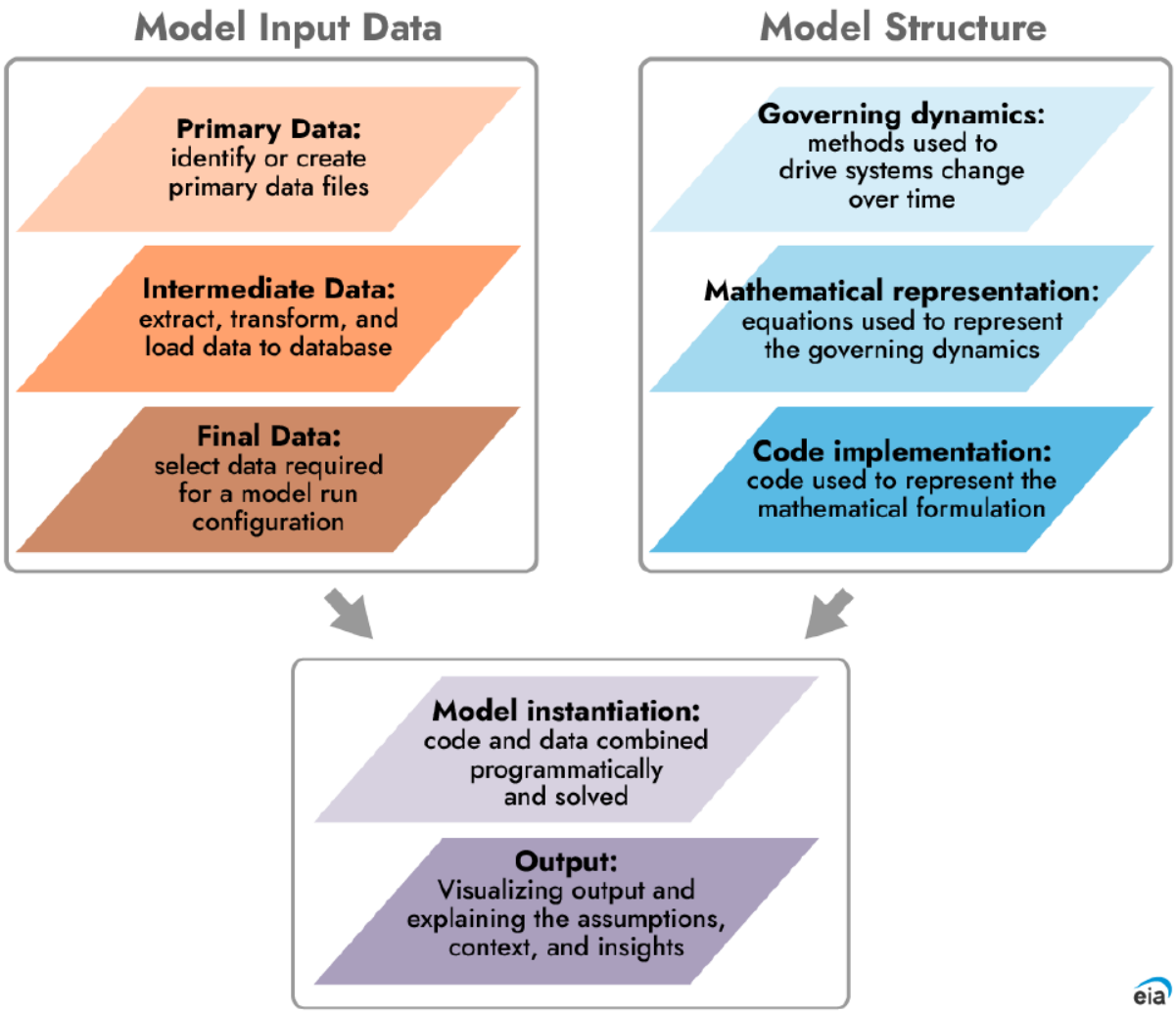


Figure 3.1 Representative layers in model development and application [123]

Models are therefore indispensable decision-support instruments. They illuminate trade-offs, stress-test assumptions, and provide shared frames of reference for diverse stakeholders. But their credibility depends not only on internal consistency and performance but also on philosophical awareness: the recognition that assumptions about truth, knowledge, and uncertainty are never purely technical. This awareness is particularly important under conditions of deep uncertainty, where competing epistemic commitments yield divergent modelling practices. Having clarified foundational concepts, the next section examines their philosophical roots (classical, modern, and contemporary), that shape how uncertainty, truth, and decision are treated in energy-system analysis.

3.3 Philosophical background: truth, uncertainty, and decision

Models are more than technical artefacts; they are structured abstractions that embody assumptions about truth, knowledge, and uncertainty. Their outputs are shaped not only by mathematics or computational design but also by deeper philosophical commitments about what can be known, how uncertainty should be represented, and how decisions ought to be justified [93]. In this sense, energy models are not neutral mirrors of reality but situated instruments that mediate between theory, data, and policy, embedding specific views of risk, uncertainty, and decision-making [117]. Accordingly, philosophical stance is not ancillary: it conditions what counts as evidence, credibility, and acceptable trade-offs in model-based reasoning.

These commitments extend beyond methods to broader epistemic positions: is the future ordered or contingent, predictable or indeterminate, objective or socially constructed? Classical distinctions between risk (measurable variability) and uncertainty (irreducible indeterminacy) remain highly influential [125], as do debates between positivist traditions, which treat models as tools of truth-discovery, and post-positivist traditions, which emphasise reflexivity, pluralism, and contingency [122]. More recent debates on wicked and super-wicked problems underscore that energy and climate challenges are not merely technical puzzles but value-laden and contested, resistant to single "optimal" solutions [86], [126].

Understanding these philosophical underpinnings is essential, as they determine not only how uncertainty is represented but also which futures are rendered plausible and legitimate within policy debates. This section therefore traces the intellectual roots of modelling, from classical antiquity through modern rationalism, probability theory, systems sciences, and late twentieth-century concepts of wickedness and resilience, showing how they continue to shape how uncertainty is framed and how decisions are supported in energy system analysis. The aim is pragmatic: to clarify why different traditions recommend different modelling choices under deep uncertainty.

3.3.1 Classical foundations: order vs contingency

The philosophical origins of thinking about uncertainty can be traced to classical antiquity. Plato regarded knowledge as the pursuit of timeless order and universal forms [127]. Reason, on this view, uncovers ideal structures that transcend everyday contingency. This Platonic orientation resonates with modern optimisation-based models, which search for least-cost or least-emission pathways as if "optimal" solutions exist independently of social or institutional context [104]. Frameworks such as MARKAL and TIMES exemplify this tendency, treating the energy system as a coherent whole and

seeking equilibrium solutions under assumptions close to perfect-foresight, that is, assuming long planning horizons with fixed expectations of technology costs, demand, and policy [128].

By contrast, Aristotle grounded knowledge in causality and empirical particulars [129]. Observation, classification, and causal explanation form the basis of understanding. This inductive orientation foreshadows bottom-up energy modelling, which emphasises technological detail, calibration, and disaggregation. Tools such as LEAP, or process-specific simulations of heat and power plants, reflect this Aristotelian view by building system behaviour from the ground up and privileging particulars over universals [130].

Other classical traditions also shaped conceptions of uncertainty. Stoic philosophy cast the cosmos as fundamentally rational, treating uncertainty as ignorance that can, in principle, be eliminated through better knowledge. This stance underpins deterministic approaches in which uncertainty is modelled as missing information to be resolved by improved data or algorithms [131], [132]. In contrast, Epicurean thought introduced clinamen, the atomic "*swerve*", as intrinsic randomness. This anticipates modern recognition of irreducible stochasticity in probabilistic analysis, Monte Carlo simulation, and stochastic optimisation [133], [134].

The enduring tension between order and contingency grounds a distinction that still shapes energy modelling; risk as calculable variability versus deep uncertainty as irreducible indeterminacy [10], [135]. Risk-oriented approaches represent uncertainty via probability distributions and address it with statistical methods, regression, or sensitivity analysis [136]. Under deep uncertainty, by contrast, modellers turn to exploratory scenarios, robust optimisation, or adaptive pathways that acknowledge the limits of prediction [9], [137]. Thus, probabilistic risk analysis may suffice for short-term fuel-price volatility (e.g., IEA, 2021), whereas disruptive dynamics, rapid electrification or unforeseen biomass constraints require exploratory scenario ensembles and robustness testing [93].

From the perspective of modelling traditions, this philosophical divide parallels the later distinction between top-down and bottom-up approaches [138]. Top-down models, emphasising macroeconomic order and optimisation, reflect a Platonic–Stoic view of a rational, largely predictable system. Bottom-up models, emphasising technological diversity and empirical disaggregation, reflect an Aristotelian–Epicurean orientation that accommodates variability and contingency. Hybrid formulations can be read as attempts to reconcile these legacies, recognising that energy systems exhibit both structural order and irreducible complexity [139], [140]. This classical lens helps explain why combining descriptive, exploratory, and prescriptive elements is often necessary in practice.

3.3.2 Early modern rationalism, inductivism, and probability

The early modern period introduced new ways of reasoning about knowledge and uncertainty that continue to shape modelling practice. René Descartes championed rationalism, emphasising deductive reasoning, methodical doubt, and the reduction of complex problems into simpler parts [141]. This methodical approach resonates with the logic of structured optimisation in modern modelling, including linear and mixed-integer programming, which reduce complex energy systems to solvable sets of constraints and decision variables.

In contrast, Francis Bacon advocated empiricism and induction, insisting that reliable knowledge arises from systematic observation and experimentation rather than pure reason [142]. This inductive orientation is echoed in statistical and data-driven methods such as regression analysis, demand forecasting, clustering, and classification, which infer patterns from empirical data rather than deductive axioms. Forecast-based energy planning models that extrapolate historical demand or technology cost trends exemplify this Baconian legacy and remain influential in national energy policy analysis.

The limitations of induction were made explicit by David Hume, who argued that past observations cannot guarantee future outcomes [143]. This "problem of induction" undermines the epistemic certainty of long-term energy demand projections, where structural breaks such as disruptive technology adoption, fuel price shocks, or sudden policy shifts can invalidate inductive forecasts. In energy system modelling, Hume's critique highlights the importance of combining data-driven projections with scenario analysis and uncertainty exploration. Together, Bacon and Hume illustrate both the power and the limits of empiricism; data can reveal trends, but it cannot guarantee their persistence under structural change.

Alongside these philosophical debates, the development of probability theory provided formal tools to treat uncertainty as quantifiable risk. Blaise Pascal and Pierre de Fermat laid the foundations of probability in their correspondence on gambling problems, while Thomas Bayes extended the framework to incorporate conditional probability and inference [144]. Pierre-Simon Laplace systematised probability theory and argued for a deterministic worldview in which, given sufficient data, the future could in principle be predicted with certainty [145]. This Laplacian determinism echoes modern "predict-and-optimise" approaches, where sufficient data and computational capability are assumed to yield reliable optimal solutions. Yet this confidence in determinism has been increasingly challenged, as contemporary transitions reveal the importance of non-linear dynamics, behavioural responses, and structural uncertainty. The upshot is methodological pluralism:

probabilistic tools remain vital, but they must be complemented by exploratory techniques when structural change is plausible.

These philosophical developments laid the groundwork for statistical modelling traditions. The formalisation of probability enabled techniques such as descriptive statistics, probability distributions, regression, Monte Carlo simulation, and later machine learning, all of which treat uncertainty as measurable variation rather than irreducible indeterminacy. In the context of energy systems, this epistemic posture is still evident in forecast-driven planning models, such as those relying on historical demand extrapolation in energy policy [93], [132]. While these approaches provide valuable insights, they also inherit the inductivist limitations identified by Hume, particularly when used to support long-term decision-making under structural uncertainty. Consequently, long-horizon planning benefits from mixing empirical induction with scenario exploration and robustness testing.

The early modern period built directly on these classical legacies, formalising their tensions through rationalism, empiricism, and the development of probability theory, foundations that continue to shape how energy models represent truth, uncertainty, and decision.

3.3.3 Nineteenth and twentieth century: formalisation and systems

The nineteenth and twentieth centuries witnessed a rapid formalisation of uncertainty into mathematical and systemic frameworks, laying the groundwork for much of contemporary modelling practice. Developments in statistical mechanics and utility theory established the idea that complex systems could be understood through aggregate regularities, while engineering prudence introduced the use of safety factors and design margins to account for variability. These practices reinforced a central conviction: that uncertainty could be disciplined and managed through quantitative representation, whether in physical systems or in decision-making contexts.

A major milestone was the creation of an axiomatic foundation for probability by Andrey Kolmogorov (1933/1950), which provided the mathematical basis for modern probability theory [146]. This enabled probabilistic modelling to be treated as a rigorous and internally consistent science. Building on this foundation, John von Neumann and Oskar Morgenstern (1944) developed expected utility theory, formalising rational decision-making under risk [147]. Leonard Savage (1954) later extended this framework with the concept of subjective probability, recognising that uncertainty can also stem from beliefs and expectations rather than only from objective frequencies [148]. Together, these advances shaped the rise of cost–benefit analysis (CBA) and a wide family of risk-based decision

frameworks, which remain influential in policy evaluation, infrastructure planning, and economic modelling.

However, the optimism that uncertainty could always be rendered calculable was tempered by Frank Knight's (1921) distinction between risk as a measurable variability and uncertainty, which is inherently immeasurable [135]. This insight continues to frame energy and climate modelling: while some factors (e.g., short-term fuel price fluctuations) can be expressed probabilistically, others (e.g., disruptive technologies, policy reversals, or social acceptance) resist quantification and demand alternative approaches.

In parallel, the systems sciences emerged as a transformative paradigm for studying complexity, interdependence, and feedbacks. Operations research (OR), developed during wartime logistics, applied optimisation methods to resource allocation and later informed the linear and mixed-integer programming (LP/MILP) techniques that remain central to energy system optimisation today [149]. Cybernetics and dynamic programming advanced understanding of control, feedback, and adaptive optimisation in complex environments. Systems thinking, formalised by Forrester (1961) [150] and extended by Meadows et al. in the influential Limits to Growth study (1972) [151], highlighted the importance of non-linear dynamics, feedback-rich processes, and long-term structural change in socio-technical systems.

At the same time, philosophers of science reframed the status of scientific knowledge. Karl Popper (1959) advanced the principle of falsifiability, arguing that science progresses through conjectures and refutations rather than through verification [152]. Thomas Kuhn (1962) introduced the concept of paradigm shifts, emphasising that scientific knowledge is provisional and subject to transformation when anomalies accumulate [153]. These perspectives resonate with energy and climate modelling, where dominant methods and assumptions are periodically overturned in response to new data, technological breakthroughs, and shifting policy priorities.

In the energy domain, these intellectual currents gave rise to powerful modelling methods that remain in widespread use. Operations research provided the foundation for LP/MILP-based optimisation models now central to power system planning, generation dispatch, and capacity expansion [154], [155], with frameworks such as MESSAGE [156], TIMES [157], and PyPSA [158]. System dynamics extended the analysis of feedback-rich processes, shaping climate–energy–economy models [150] such as World3 [159] and later integrated assessment models (IAMs) [160], as well as contemporary digital twins used for real-time simulation and control [161]. In parallel, probabilistic frameworks such as Monte Carlo simulation [133] and stochastic optimisation [162] became standard for assessing uncertainty in resource planning, reliability analysis, and investment appraisal.

Collectively, these developments signalled the maturation of modelling as both a science and a decision-support practice. They established many of the quantitative and systemic traditions that underpin energy system analysis today, while also reinforcing Knight's enduring insight that not all uncertainty is reducible to risk. This dual legacy, formal probabilistic treatment alongside systems thinking, sets the stage for modern deep uncertainty handling methods.

3.3.4 Late 20th century: bounded rationality, wicked problems, and resilience

The late twentieth century marked a decisive shift in how uncertainty, complexity, and decision-making were conceptualised. Building on Frank Knight's earlier distinction between risk (measurable variability) and uncertainty (immeasurable indeterminacy), Herbert Simon's theory of bounded rationality (1957) challenged the assumption that decision-makers pursue global optima. Instead, he argued that choices are constrained by limited information, cognitive capacity, and time, leading actors to adopt "*satisficing*" strategies, solutions that are good enough rather than perfectly optimal [163]. This perspective foreshadowed the use of heuristic and metaheuristic methods in energy planning, where exact optimisation is often computationally intractable or misaligned with real-world decision-making. Techniques such as genetic algorithms, simulated annealing, and other search heuristics reflect this pragmatic orientation toward feasible rather than perfectly optimal solutions [164], [165].

At the same time, Horst Rittel and Melvin Webber (1973) introduced the concept of *wicked problems*, characterised by contested objectives, shifting framings, and the absence of definitive solutions [24]. Their insights deeply influenced policy analysis and planning, emphasising that societal challenges such as climate change or industrial decarbonisation are not amenable to purely technical fixes. In the modelling domain, wicked problems spurred the development of multi-criteria decision analysis (MCDA) and participatory approaches that explicitly integrate diverse values, objectives, and stakeholder perspectives [166]. More recently, scholars have described climate change as a "*super wicked problem*" due to its urgency, self-reinforcing dynamics, and institutional inertia [86]. Together, these framings underscore that energy transition pathways cannot be resolved through optimisation alone but require frameworks that accommodate plurality, contestation, and adaptation.

In ecology, C. S. Holling (1973) advanced the concept of resilience, defining it as the capacity of a system to absorb shocks and disturbances while maintaining its essential functions [102]. This view shifted attention from stability and optimisation toward adaptability and persistence under stress. In energy system analysis, resilience metrics have increasingly been incorporated into assessments of energy security, infrastructure reliability, and the robustness of supply chains [167], [168].

These intellectual shifts also found practical expression in the rise of scenario planning, most prominently through Shell's strategic foresight exercises in the 1970s [99]. Scenario planning, unlike probabilistic forecasting, does not seek to predict a single future but instead explores multiple plausible futures, enabling organisations to test strategies across diverse contexts. When combined with participatory MCDA, such approaches reflected the growing recognition that energy and climate decisions are value-laden, contested, and deeply uncertain. These contributions collectively reoriented modelling from prediction towards learning and adaptation.

In energy system modelling, these contributions translated into new practices that complemented traditional optimisation and systems analysis. MCDA has become widely used in technology evaluation, for example, in comparing the trade-offs of boilers, heat pumps, and biomass systems under multiple criteria [169]. Scenario planning has been applied in the New Zealand Emissions Trading Scheme (NZETS) and the GIDI Fund assessments, helping policymakers consider divergent futures rather than a single reference pathway. Resilience analysis has gained traction in evaluating the robustness of electricity grids, biomass supply chains, and process heat transitions, particularly in the face of climate risks and infrastructure disruptions.

Collectively, these approaches reinforced the importance of pluralism, adaptability, and stakeholder engagement in energy modelling. They prepared the ground for the twenty-first century paradigms that explicitly address deep uncertainty, shifting the emphasis from prediction and optimisation to robustness, flexibility, and legitimacy in decision-making.

3.3.5 Deep uncertainty and the decision turn

Uncertainty has long been a central concern in both philosophy and decision sciences. Classical accounts distinguished between situations of risk, where probabilities can be assigned to outcomes, and uncertainty, where probabilities are unknown or unknowable. Frank Knight's seminal work formalised this distinction, noting that risk refers to quantifiable variability, while uncertainty arises when the distribution of future states cannot be specified in probabilistic terms [135]. In the context of energy systems, risks may include short-term fluctuations in fossil fuel prices with well-defined historical distributions, whereas uncertainty encompasses structural shifts such as disruptive technology adoption, unforeseen policy reversals, or social contestation [107], [137].

Philosophical traditions have offered multiple orientations toward uncertainty. Rationalist perspectives, from Descartes to contemporary optimisation models, treat uncertainty as a deficiency of information to be reduced through reason, logic, and data [141], [170]. Empiricist traditions, emphasised by Bacon [142] and later critiqued by Hume, acknowledge the provisional nature of

knowledge, recognising that past observations cannot guarantee future outcomes, a problem that undermines long-term extrapolation in energy demand forecasts [171]. Pragmatist views, associated with Dewey and later system dynamics, stress usefulness over prediction; models are valued not for their correspondence to a single "*true*" future but for their ability to reveal patterns and inform adaptive strategies [150], [172]. More recently, constructivist and post-normal science perspectives frame uncertainty as socially embedded and irreducible, requiring plural approaches, stakeholder participation, and reflexivity in modelling [107], [173].

This conceptual evolution has led to the modern category of deep uncertainty, defined as conditions where stakeholders cannot agree on (i) the appropriate models, (ii) the probability distributions of key parameters, or (iii) the metrics by which outcomes should be judged [9], [10]. Deep uncertainty is highly characteristic of industrial decarbonisation; the pace of electrification, the resilience of biomass supply chains, and the trajectory of carbon pricing are all deeply contested and resistant to probabilistic specification. Traditional optimisation or equilibrium approaches, which assume predictable structures and calculable risks, struggle under such conditions, often producing misleading precision or inconsistent results across models [93], [107].

By the turn of the twenty-first century, the fragility of the predict-and-plan paradigm had become increasingly apparent. Forecast-based strategies rely on stability in trends and consensus on objectives, yet real-world experience shows that these assumptions frequently break down. Volatility in emissions trading schemes, the financial risks of stranded assets, and abrupt policy reversals illustrate the limits of planning based on single best-estimate forecasts [19], [137].

In response, decision science advanced a suite of methods under the umbrella of Decision Making under Deep Uncertainty (DMDU) [9]. These approaches share a common philosophy: instead of "*predict-and-optimise*", they embrace "*explore-and-adapt*". Robust Decision Making (RDM) emphasises strategies that perform satisfactorily across a wide ensemble of plausible futures [10]. Dynamic Adaptive Policy Pathways (DAPP) structures decisions as adaptive sequences that can be revised as conditions evolve [50]. Exploratory Modelling and Analysis (EMA) systematically explores uncertainty through ensembles of model experiments [19], [174]. Other variants include info-gap decision theory [175] and adaptive robust optimisation [176], all prioritising robustness and flexibility over predictive certainty.

These approaches have gained significant traction in energy and climate modelling, particularly where long-term planning is exposed to technological volatility and political contestation. EMA and RDM have been applied to climate mitigation portfolios and capacity expansion planning [177]. In New Zealand, frameworks such as the Regional Energy Transition Accelerator (RETA) have begun

experimenting with adaptive planning approaches that explicitly acknowledge uncertainty in biomass supply and electrification pathways [16]. Similar approaches are emerging in process heat decarbonisation, where the goal is not to identify a single least-cost pathway but to ensure strategies remain viable under multiple plausible futures. In short, DMDU reframes modelling success from accurate prediction to durable performance under change.

Collectively, these methods mark a decision turn in modelling practice. They reject the belief that uncertainty can be fully eliminated through better forecasts and instead embrace plurality, contingency, and resilience. Models are increasingly seen less as forecasting machines and more as decision-support laboratories, designed to stress-test strategies under conditions of deep uncertainty. This orientation directly informs the modular, site-to-region framework developed in this thesis, which combines deterministic techno-economic optimisation with ensemble-based exploratory analysis to support robust, least-regret planning for industrial process heat decarbonisation.

3.3.6 Epistemic beliefs and modelling practice

The trajectory of energy modelling has been shaped not only by methodological innovations but also by deeper epistemic beliefs about what models represent and how they should be used. As Bock and Pfenninger-Lee (2025) argue [120], modellers often operate within distinct epistemic orientations that can be broadly described as positivist and postpositivist.

Positivist traditions treat models as objective instruments of truth-discovery. In this view, models are designed to capture reality as faithfully as possible, abstracting complexity into coherent representations of system dynamics. Their legitimacy derives from internal rigour, formal consistency, and predictive capacity. Large-scale optimisation frameworks such as TIMES-NZ [178] or global Integrated Assessment Models (IAMs) (e.g., MESSAGE, IMAGE, REMIND) exemplify this orientation. By employing linear programming, least-cost optimisation, and long-term forecasting, these models aim to provide policy-relevant insights grounded in the premise of scientific objectivity [179].

Post-positivist approaches, in contrast, emphasise the socially embedded and contextual nature of modelling. Here, models are seen not as neutral mirrors of reality but as situated constructs that embody values, institutional framings, and normative choices. Their legitimacy rests less on claims of objectivity and more on transparency, inclusivity, and reflexivity. Examples include exploratory scenario modelling [107], participatory modelling workshops [93], [159], and open-source initiatives such as the Open Energy Modelling (Openmod) community [180]. These practices increase contestability and reuse but can be perceived as less prescriptive. Bridging these orientations calls for

systems-aware practice; clear interfaces, documented assumptions, and audit trails that make models both tractable and contestable.

This epistemic divide has important implications for modelling practice. Deterministic optimisation models often gain political traction because their results appear authoritative and definitive; yet this very authority can obscure the contingent assumptions on which they rest, such as discount rates, fuel prices, or behavioural adoption pathways [93], [181]. By contrast, exploratory and participatory models make assumptions explicit and foreground uncertainty, but they are sometimes dismissed as less rigorous or overly open-ended [182], [183]. The credibility of scenarios and pathways therefore depends not only on their technical quality but also on how epistemic positions are negotiated and communicated within the modelling process [173]. Bridging these orientations calls for systems-aware practice: clear interfaces, documented assumptions, and audit trails that make models both tractable and contestable.

3.4 Systems definition and systems thinking

At its most general, a system may be defined as a set of objects together with the relationships between them and between their attributes [184]. This definition emphasises that a system is more than the sum of its parts: it is constituted by the relations that bind components together into an organised whole. For energy systems, this means that boilers, turbines, transmission lines, markets, and regulations form a coherent system not merely by co-existence but by their structured interactions, dependencies, and functions. It also directs attention to processes (flows, controls, delays) and to the institutional rules that shape those interactions.

Energy systems are therefore best understood as socio-technical constructs, combining material flows of fuels, heat, and electricity with institutional frameworks, behavioural routines, and policy regimes [150]. Transitions in such systems are not reducible to technological substitution alone and reflect co-evolution between infrastructures, governance, and cultural practices, which can create forms of lock-in and resistance to change. The Multi-Level Perspective (MLP) further elaborates this by distinguishing between niche innovations, stabilised socio-technical regimes, and broader landscape pressures, showing how feedbacks across these levels shape both continuity and disruption [184]. For industrial process heat, this implies that equipment choices, market incentives, and regulatory settings co-evolve, constraining what is feasible at any point in time.

Systems thinking provides a conceptual toolkit for capturing these properties. Building on general systems theory [150] and system dynamics [185], it highlights feedback, interdependence, and adaptation as defining characteristics of complex socio-technical systems. Such systems display

nonlinearity, path dependence, and emergent behaviour, meaning outcomes cannot be fully anticipated by analysing components in isolation [186]. In practice, interventions frequently generate unintended consequences or policy resistance, as reinforcing and balancing feedback loops alter the intended direction of change [187]. Path dependency is especially salient in energy; once capital investments, institutions, and knowledge trajectories align along a given course, they constrain alternative pathways and bias decisions toward continuity [188]. Time lags (e.g., connection queues, consenting lead times) and irreversibility (e.g., sunk costs, contractual lock-in) further condition transition pace and sequencing. Where hourly chronology is material, but full chronology is computationally burdensome, representative-time methods (e.g., representative days or hours) can reduce dimensionality while preserving key statistics relevant to operations, enabling tractable many-run exploration.

A further contribution of systems thinking lies in its explicit attention to hierarchy and scale [189]. Energy systems comprise nested subsystems that interact across local, regional, national, and global levels. In process heat decarbonisation, site-level processes and utilities (steam, hot water, refrigeration) and on-site electricity demand operate within GXP (Grid Exit Point) and sub-regional contexts (e.g., coal and biomass suppliers) shaped by network constraints and fuel logistics, themselves embedded in regional resource endowments and national policy settings [186], [189]. Accordingly, a site → GXP/sub-region → region architecture with explicit interfaces and 'thin-waist' data products such as hourly demand profiles, capacity envelopes, and price (and constraint) signals supports iterative, bidirectional alignment between local options and system limits.

Recent literature on energy regions confirms this multi-level character, showing how decentralised initiatives and local autonomy efforts can accumulate into distinctive regional trajectories [190]. Similarly, research on regional integrated energy systems demonstrates that effective decarbonisation requires accounting for both technical couplings across electricity, heat, and fuel systems and the governance arrangements that mediate decision-making across scales [191]. Work on decentralised energy autonomy further illustrates how bottom-up initiatives and municipal-scale projects introduce distinctive forms of resilience and flexibility, while still depending on broader system integration [192]. A modular, site-to-region perspective therefore reflects the inherent organisation of energy systems; it accommodates micro-level heterogeneity while recognising the interdependencies that structure meso- and macro-scale outcomes. In operational terms, this supports iterative ("co-simulation") passes between site models and regional modules to resolve feasibility under changing constraints.

For this thesis, systems thinking offers both a conceptual lens and a methodological foundation. Conceptually, it reframes industrial decarbonisation as a complex transformation shaped by

interdependent socio-technical feedbacks, rather than as a linear process of technology substitution. Methodologically, it motivates modelling approaches capable of representing dynamic interactions, cross-scale couplings, and emergent properties while retaining transparency and modularity. This perspective directly informs the design principles set out in following chapters and the site-to-region modelling architecture developed subsequently, where explicit interfaces, hourly resolution (or representative days), and reproducible data contracts are used to maintain both tractability and decision visibility.

3.5 Principles for next-generation modelling

The preceding discussion highlights that energy systems are socio-technical, multi-level, and deeply uncertain. To capture these characteristics, next-generation modelling frameworks must embody several key principles:

Modularity and Scalability

Energy systems are organised in nested hierarchies, from site-level processes to regional and national infrastructures. Models must mirror this organisation by being modular in design; individual components, such as technologies, processes, or regions, should be independently representable yet interoperable within a larger framework via explicit interfaces and “thin-waist” data products. Modularity allows fine-grained site-level representations to inform regional balances while supporting aggregation to national or international analyses. Scalability ensures that models can evolve from targeted studies of specific sites to comprehensive evaluations of system-wide transitions.

Transparency and Openness

Credibility depends on clarity. Models should make explicit the assumptions, boundary conditions, and data that underpin their results. This not only avoids the opacity of “black box” modelling but also strengthens reproducibility and trust. Use version-controlled configurations, input/output provenance, and openly documented schemas. Openness further extends to data and methods, where transparent documentation and shared standards enable broader scrutiny and collaborative development, making models more resilient to critique and more relevant for decision-making.

Flexibility and Adaptability

Transition pathways are dynamic, shaped by technological innovation, behavioural change, and evolving policy landscapes. Models must therefore be designed with flexibility and should be able to incorporate new data sources, emerging technologies, and evolving sectoral linkages through

configuration rather than code refactoring. Adaptability ensures that frameworks remain decision-relevant over time, even as conditions shift.

Robustness and Resilience

Complex systems are subject to shocks, feedbacks, and path-dependent lock-ins. Next-generation models should therefore prioritise outputs that remain meaningful under a range of assumptions, rather than optimised only for a single expected future. Report decision-visible metrics (e.g., reliability, satisficing shares, regret distributions) alongside cost and emissions. Robustness reflects the need for models to generate insights that hold across diverse conditions, while resilience captures their ability to assess system performance under stressors such as supply disruptions or climate extremes.

Granularity and Data Integration

Energy systems, particularly in industrial decarbonisation, require attention to micro-level detail. Models must capture site-specific processes, operational variability, and regional resource constraints, at hourly or representative-day resolution where material, while ensuring that this granularity scales upward to inform broader analyses. Integrating process-level and regional data allows for more accurate reflection of heterogeneity and for uncertainty at the local level to be systematically incorporated into higher-level transition assessments.

Pluralism and Reflexivity

Transitions are not solely technical but contested, involving economic trade-offs, institutional dynamics, and social/cultural values. Models must therefore accommodate multiple perspectives and framings. Pluralism acknowledges that no single lens can capture the complexity of energy transitions, while reflexivity requires modellers to be explicit about how assumptions, values, and worldviews shape the construction of models. Assumption logs and stakeholder-visible levers increase both legitimacy and relevance.

Interoperability and Ethical Orientation

Energy systems increasingly cut across sectoral and disciplinary boundaries. Models must therefore be interoperable, with open standards that allow data and outputs to be exchanged across platforms. Where electricity-system detail is material, interoperate with established open electricity models rather than duplicating functionality. Beyond technical coupling, ethical considerations should also guide design. Models should integrate concerns for distributional effects, representation, and justice, ensuring that transitions are evaluated not only in terms of efficiency, cost or emissions but also in terms of fairness and equity.

Taken together, these principles define next-generation energy system models as decision-support laboratories that extend beyond traditional forecasting machines. Such models retain the ability to generate forecasts and outlooks, but their distinctive value lies in exploring uncertainty, testing strategies, and revealing trade-offs. They combine modularity with scalability, transparency with openness, and robustness with reflexivity, while embedding granularity, interoperability, and ethical orientation. In doing so, they produce decision-visible artefacts (metrics, signposts, and triggers) that are auditable and policy-ready. These principles represent the conceptual design requirements for future models, and the next step is to synthesise what they imply for transition and decision support before operationalising them in the methodological architecture.

3.6 Synthesis and implications for transition and decision support

This section has framed industrial decarbonisation, particularly process heat, as a decision problem characterised by heterogeneity, path dependence, and deep uncertainty. Three implications follow for how transition strategy and decision support should be structured.

First, decision support must be explicitly multi-scale and interface-based. Credible transition guidance cannot remain confined to either site engineering feasibility or system-level envelopes in isolation. Instead, it should treat cross-scale interaction as a first-class design requirement: site choices create temporally structured demands on shared infrastructures and resources, while regional constraints (capacity, delivery, and institutional frictions) delimit which site pathways remain credible. Practically, this implies the use of explicit interface variables and stable exchange contracts that allow local realism and system constraints to be evaluated together without collapsing into opaque, monolithic representations.

Second, deep uncertainty shifts the evaluative focus from optimality to vulnerability and robustness. Where key drivers cannot be credibly assigned stable probabilities, decision support should prioritise stress-testing strategies over ensembles of plausible conditions, characterising where and why strategies fail (or become unacceptable), and identifying forms of flexibility that reduce regret and tail exposure. This implication is inherently sequential: it encourages adaptive reasoning (using signposts and triggers) because robustness is not only about “picking once” but also about maintaining decision agility as conditions evolve.

Third, legitimacy in high-stakes transitions requires that transparency and reproducibility be treated as validity conditions, not optional presentation choices. In settings with sparse or aggregated evidence, the credibility of conclusions depends on whether assumptions, transformations, and evaluation steps remain inspectable and reproducible. Interoperability further strengthens this

legitimacy by enabling appropriate coupling to established models where specialised detail is material, while keeping the overall decision narrative auditable and comprehensible to non-specialist stakeholders.

The conceptual foundations developed here imply three design commitments for decision-centred modelling, which Chapter 6 carries into the DCM architecture:

- Deep uncertainty to ensembles plus vulnerability characterisation, not single best-estimate optimisation. Robustness should be demonstrated through performance across many plausible conditions, with explicit identification of failure regions and trade-offs.
- Epistemic humility to governed reproducibility is part of validity. Audit trails, documented assumptions, versioned configurations, and acceptance criteria are methodological safeguards that protect interpretability and stakeholder trust under data limitations.
- Decision relevance to separate a deterministic backbone from exploratory evaluation. A stable feasibility core supports interpretability and comparability, while exploratory layers (ensembles and diagnostics) reveal sensitivity, tail risk, and conditions under which preferences reverse.

Together, these implications establish the evaluative yardstick used in the remainder of the thesis: models are assessed not only by representational detail, but by their ability to support auditable, cross-scale, uncertainty-aware decision support. The next chapter therefore reviews major modelling paradigms through this lens, identifying the approaches most capable of meeting these requirements for industrial process-heat transitions.

Chapter 4
Modelling Paradigms for Industrial and Energy Systems: Strengths, Limits, and Decision-Scale Gaps

4.1 Introduction

Moving from principles to practice, the design of energy system models is guided by underlying paradigms and approaches that shape what is emphasised, how complexity is represented, and which forms of decision support are enabled. These paradigms provide the interpretive lens through which energy transitions are analysed and understood. In this context, a paradigm in energy modelling can be defined as a structured way of framing a problem, selecting assumptions, and organising methods to generate analytical insights. Paradigms establish the boundaries of what is represented in a model, how uncertainty is addressed, and which decisions the model is intended to inform. They are therefore not only technical constructs but also epistemic commitments, translating philosophical orientations into practical modelling strategies [93], [122].

The literature commonly identifies three structural paradigms: top-down, bottom-up, and hybrid approaches. Top-down models situate energy within an economy-wide framework, often using computable general equilibrium (CGE) or econometric methods to examine how energy interacts with macroeconomic dynamics such as trade, consumption, and gross domestic product [193], [194]. Bottom-up models take the opposite perspective, representing energy systems through detailed technology characterisations, engineering constraints, and process-level data. Hybrid approaches integrate the two perspectives, either by coupling separate models or by embedding feedback loops that capture both economy-wide drivers and technological feasibility [11], [93].

Beyond this classical tripartite division, models can also be classified by their methodological orientation (e.g., optimisation, simulation, equilibrium) [93], [195] and their decision-support role (e.g., descriptive, prescriptive, exploratory, or robust) [174], [196]. As summarised in Table 4.1, these categories represent complementary perspectives on how energy system models can be understood, and they have been widely employed in reviews and comparative studies of energy system modelling [93], [193], [194].

Table 4-1 Classification lenses in energy system modelling

Classification Lens	Categories	Core Idea / Purpose
Structural	Top-down, Bottom-up, Hybrid	Defines system boundary: economy-wide, technology-detailed, or integrated
Methodological	Optimisation, Simulation, Equilibrium	Distinguishes models by solution approach: least-cost vs dynamic/behavioural vs market-clearing
Decision Orientation	Descriptive, Prescriptive, Exploratory, Backcasting, Robust/Adaptive	Distinguishes models by decision-support role: documenting, prescribing, envisioning long-term goals, exploring uncertainty, or stress-testing for resilience

In addition to paradigms, the literature highlights a set of extended classification dimensions that strongly influence how models are applied and interpreted. These are not paradigms in the strict sense but rather design features that condition the scope, tractability, and applicability of modelling exercises [197], particularly visible in the context of process heat decarbonisation. They include spatial resolution, which is especially critical for industrial process heat, where geographically concentrated facilities must be analysed within site-to-regional-to-national hierarchies [198]; sectoral scope, since process heat cannot be separated from electricity and biomass supply without distorting substitution and competition effects [11]; and temporal resolution, given the importance of hourly and seasonal variability for both process heat demand and integration with the electricity grid [199]. Openness and transparency are equally vital, shaping reproducibility, stakeholder trust, and policy legitimacy in contested decarbonisation debates [200].

Other dimensions, such as market behaviour and data intensity, remain relevant but are of secondary importance for process heat modelling. Market competition is central to electricity sector models but less so for industrial users, whose choices are more strongly shaped by technology adoption decisions [195]. Similarly, data intensity requires balancing the benefits of detailed site-level information with computational tractability, motivating modular and hierarchical approaches [193]. Emerging paradigms such as artificial intelligence, machine learning, and digital knowledge systems can also be understood as extended dimensions: they do not replace classical paradigms, but they reshape how data are handled, how scenarios are generated, and how models interact with decision processes [161], [201], [202].

Taken together, these lenses demonstrate that energy system modelling is not defined by a single dominant paradigm but by a spectrum of approaches and design dimensions. For process heat, the combination of site-specific representation, sectoral coupling, temporal detail, and open modular design emerges as especially significant, offering a pathway to manage deep uncertainty in fuel supply

and policy while supporting resilient planning. The following sections examine these paradigms and dimensions in more detail, with particular attention to their implications for industrial decarbonisation in New Zealand.

4.2 Structural classification

Structural paradigms represent one of the most established ways of classifying energy system models, defining the boundaries within which energy supply, demand, and policy interactions are represented. From their origins in the 1970s oil crises, when energy–economy linkages and long-term planning became urgent, these paradigms have provided contrasting but complementary lenses for capturing the complexity of modern energy systems [93], [193]. At their core, structural paradigms distinguish whether emphasis is placed on economy-wide interactions, detailed technological and engineering processes, or an integration of both perspectives. Top-down approaches draw primarily from economics and general equilibrium theory, highlighting market and price responses; bottom-up models stem from engineering traditions, focusing on process-level technology representation and optimisation; while hybrid approaches attempt to reconcile the two, linking macroeconomic drivers with technology-rich system descriptions [195]. Together, these paradigms continue to underpin contemporary modelling practice internationally, shaping not only methodological choices but also the type of insights that can be generated for energy policy, planning, and transition analysis, particularly under deep uncertainty, where robust hybrid designs are increasingly vital [9].

4.2.1 Top-down approaches

Top-down models embed energy within the broader macroeconomic system, treating it as one sector among many that interact through markets, prices, and aggregate demand. They draw heavily on economic theory and have become a central tool for assessing the economy-wide consequences of energy and climate policy. Two dominant forms can be distinguished: computable general equilibrium (CGE) models and econometric/macro econometric models.

CGE frameworks apply microeconomic foundations and equilibrium assumptions to simulate the entire economy. Producers are typically modelled as profit maximisers and households as utility maximisers, interacting through markets that clear simultaneously. Energy appears as a production input alongside capital and labour, with substitution possibilities determined by constant elasticity of substitution (CES) functions. This structure allows the analysis of how shocks, such as carbon pricing, fuel taxes, or productivity changes, reverberate through trade, income distribution, and gross domestic

product. The Global Trade Analysis Project energy model (GTAP-E) is a widely used global CGE extension that incorporates energy substitution and emissions trading mechanisms [203]. Applications include assessments of carbon taxes, biofuel mandates, and international climate agreements. However, CGE models face persistent criticisms: parameters (e.g., substitution elasticities) are often weakly validated, results may be sensitive to calibration choices, and the assumption of perfectly competitive equilibria can limit realism [204].

At the national level, many treasuries and policy agencies have adopted CGE models to simulate domestic economic and policy scenarios. In New Zealand, the Treasury developed the New Zealand Treasury Model (NZTM), a dynamic CGE framework designed to represent the small open economy, incorporating production sectors, households, government, and external trade. It has been used to evaluate fiscal reforms, external price shocks, and environmental policies [205]. More recently, the C-PLAN model (Climate PoLicy ANalysis) has been applied for climate policy evaluation, with a particular focus on emissions mitigation in agriculture, forestry, and industrial sectors. Developed with support from the Climate Change Commission, C-PLAN provides recursive dynamic simulations of climate and energy policy impacts across the New Zealand economy [206]. Other notable examples include the BERL “Joanna” model, which disaggregates more than 50 industries and 40 occupational categories for applied trade and labour market analysis [207], and a multi-regional CGE model developed at Victoria University of Wellington to capture spatial economic interdependencies and the distribution of shocks across regions [208].

Despite their theoretical rigour, top-down models typically assume perfectly competitive markets and representative agents, limiting the treatment of market imperfections, heterogeneity, and bounded rationality [209]. Outcomes are sensitive to functional forms (e.g., CES) and parameterisation, particularly substitution elasticities between energy, labour, and capital [210]. These elasticities are often drawn from global datasets or calibrated rather than empirically validated, raising concerns about robustness. Moreover, top-down frameworks usually operate at an aggregated sectoral level, meaning technological detail is stylised and industrial processes are broadly categorised rather than represented with engineering specificity [211].

This makes them less suitable for analysing site-level or technology-specific transitions such as the decarbonisation of process heat at facilities like Edendale dairy plant, where grid constraints and biomass logistics defy aggregate assumptions, unless complemented by bottom-up or hybrid models. Critical perspectives have further argued that relying solely on CGE for sustainability or transition assessments risks overlooking complexity, path dependency, and structural change [212].

4.2.2 Bottom-up approaches

While top-down models prioritize economy-wide feedbacks, bottom-up models characterise energy systems from the level of end-use technologies upward, linking service demand to technology performance, process constraints, and resource availability [213]. Their defining feature is the explicit representation of technologies, fuel chains, and engineering feasibility, which enables detailed assessment of efficiency improvements, fuel substitution, and system restructuring. In contrast to top-down models, which rely on aggregate production functions, bottom-up approaches provide technology-rich pathways that can capture resource competition and operational constraints with greater granularity [104].

The dominant stream of bottom-up modelling applies optimisation techniques, typically linear or mixed-integer programming, to identify least-cost energy system configurations subject to resource, technological, and policy constraints. Canonical examples include the MARKAL and TIMES families, developed under the IEA-ETSAP programme, which integrate technology databases with investment, operation, and retirement decisions across multi-decadal horizons [214], [215]. Their integrated assessment variant, TIAM, links energy transitions with climate outcomes in global scenarios. MESSAGE and its successor MESSAGEix, developed at IIASA, provide similar optimisation capabilities while emphasising scenario flexibility and integration with integrated assessment frameworks [93]. More recently, open-source models such as OSeMOSYS, TEMOA, and Calliope have advanced transparency, reproducibility, and ensemble-style analysis [216], aligning with the broader trend toward open, modular, and uncertainty-aware tools. Collectively, these frameworks have underpinned global mitigation assessments (e.g., IPCC scenarios) and national planning studies, reinforcing their central role in long-term transition analysis.

A parallel branch of bottom-up approaches emphasises simulation and accounting frameworks, which prioritise transparency and accessibility over optimisation rigour. The Long-range Energy Alternatives Planning system (LEAP) is the most widely used, allowing stakeholders to construct demand–supply–emissions scenarios without solving optimisation problems [130]. Its accessibility has made it particularly valuable in capacity-building and policy support in developing economies [213]. EnergyPLAN provides higher temporal granularity, modelling hourly operational feasibility of high-renewable systems [217]. In electricity contexts, open toolkits such as PyPSA complement bottom-up energy models by adding power-flow and operational realism, bridging long-term planning with short-term dispatch analysis [158].

In New Zealand, bottom-up approaches have played a prominent role in energy planning and policy analysis. The TIMES-NZ model, initially developed by the Business NZ Energy Council and now

extended by EECA, represents the national energy system from primary resources through to end-use demand. TIMES-NZ 2.0 has informed national scenario exercises such as the BEC 2050 “Kayak” and “Waka” pathways and continues to shape policy debate [111], [178], though critiques highlight TIMES’ deterministic assumptions limiting deep uncertainty handling [218]. At the regional level, EECA’s Regional Energy Transition Accelerator (RETA) initiative applies detailed engineering characterisation to industrial clusters, mapping opportunities for electrification, biomass switching, and efficiency improvements in process heat [16]. These applications highlight how bottom-up models offer the granularity required to evaluate site-specific and sectoral transitions that aggregated top-down models cannot capture.

The strengths of bottom-up approaches lie in their explicit representation of technologies, direct mapping of policies to technology adoption, and capacity to test system feasibility under physical and engineering constraints [104]. These attributes make them indispensable for industrial decarbonisation, where site-level heterogeneity, temperature requirements, and utility integration strongly influence technology choices. However, they also face well-documented limitations. Classical formulations often assume perfect foresight and deterministic optimisation, limiting their ability to capture uncertainty, behavioural diversity, or path dependency [213]. Results are highly sensitive to temporal aggregation: representative time slices may distort the role of storage or bias renewable integration [93]. Data requirements are particularly intensive, especially at the site level where process heat data have historically been scarce, a gap only recently addressed through EECA’s Regional Heat Demand Database (RHDD) [7]. Simulation-oriented frameworks, while transparent and stakeholder-friendly, rely on exogenous assumptions and lack the optimisation depth needed for technology pathway evaluation.

Taken together, bottom-up approaches provide the necessary granularity for evaluating process heat decarbonisation pathways in New Zealand and internationally. Yet their limitations underscore the need for complementarity: integration with macroeconomic models to capture economy-wide effects, coupling with power-system tools to assess grid impacts, and embedding robust decision-making methods to address uncertainty. These extensions are particularly relevant for next-generation frameworks that aim to move beyond deterministic foresight and provide robust, stakeholder-relevant insights into industrial transitions.

4.2.3 Hybrid approaches

Hybrid approaches integrate the strengths of top-down and bottom-up paradigms to provide a comprehensive framework for analysing energy–economy–environment interactions. By combining

macroeconomic dynamics with detailed technological characterisation, these models capture the complex interplay driving energy transitions [104], [211]. Integrated Assessment Models (IAMs), such as MESSAGE, TIMES, and REMIND, are prominent examples of hybrid frameworks, widely used in global climate policy (e.g., IPCC AR6 scenarios) to evaluate how carbon policies shape economic growth, technology deployment, and emissions pathways [215], [219], [220]. IAMs bridge socioeconomic systems (e.g., GDP, population growth) with energy and climate modules, offering a holistic view of mitigation strategies under uncertainty [97]. Two principal strategies for hybridisation are distinguished in the literature: *soft-linking* and *hard-linking* [215].

Soft-linking connects independent top-down and bottom-up models through iterative exchange. For instance, a CGE model might project economy-wide energy demand, which a bottom-up TIMES model uses to optimise technology choices (e.g., solar vs. coal). The TIMES model then updates costs or emissions, feeding back to the CGE for recalibration [221]. This approach is flexible, as seen in IPCC scenarios where MESSAGE outputs inform macroeconomic models, but requires careful alignment to ensure consistency across assumptions (e.g., fuel prices, carbon taxes) [211]. Misaligned boundaries can lead to divergent results, limiting reliability [222].

Hard-linking embeds both macroeconomic and engineering components in a single framework, solving equilibrium and optimisation constraints simultaneously. Examples include TIMES-MACRO and MESSAGE-MACRO, where a macroeconomic growth module (e.g., GDP-driven energy demand) is directly coupled with a technology-rich energy system (e.g., renewable penetration rates) [156]. This ensures consistency but increases computational complexity, often requiring high-performance computing, and reduces transparency due to intricate model structures [222]. For example, MESSAGE-MACRO's integration of global trade and energy flows supports IPCC's 1.5°C pathways but can obscure specific technology assumptions for non-expert users [219].

Hybrid models, particularly IAMs, are pivotal in global and national energy planning. In IPCC assessments, MESSAGE and REMIND analyse trade-offs between economic growth and low-carbon technology adoption, informing Paris Agreement targets [219], [220]. Nationally, hybrid frameworks assess industrial decarbonisation (e.g., steel, cement) alongside electricity market dynamics and macroeconomic adjustments, offering nuanced insights compared to single-paradigm models [211]. In New Zealand, hybrid approaches are emerging, though not yet fully hard-linked. The TIMES-NZ model, used by MBIE for emissions pathways, has been soft-linked with CGE models such as NZTM and C-PLAN to align technology choices with economy-wide impacts [223]. EECA's process heat studies, for example, compare TIMES-NZ outputs with CGE projections to evaluate biomass constraints under volatile ETS prices, ensuring policy coherence [38], [46].

The strength of hybrid approaches lies in their ability to balance macroeconomic credibility with technological realism [211]. They reduce the risk of overlooking either aggregate system feedbacks or engineering constraints, providing a more comprehensive platform for policy evaluation, including robust stress-testing of strategies across uncertain futures [137]. However, they inherit limitations from both paradigms, including high data demands, computational complexity, and challenges in aligning diverse assumptions [221]. Transparency is a key issue; merging aggregated economic models with detailed engineering systems can create a "black box" effect, hindering stakeholder trust [93]. For instance, IAMs' long-term projections (e.g., to 2100) often oversimplify human behaviour or technological breakthroughs, risking overly optimistic or pessimistic scenarios [221]. The distinctions across top-down, bottom-up, and hybrid paradigms are summarised in Table 4.2, which highlights their core logics, strengths, limitations, and representative models.

Table 4-2 Comparison of structural paradigms in energy system modelling

Paradigm	Core Logic	Strengths	Limitations	Representative Models
Top-down	Economy-wide macroeconomic representation	Captures trade, income, and fiscal interactions; consistent with policy shocks	Limited technology detail; assumes equilibrium and representative agents	GTAP-E, NZTM, C-PLAN
Bottom-up	Technology-rich, engineering-based representation	Granular technology detail; explicit resource and process constraints	High data needs; deterministic foresight; less suited to behavioural dynamics	TIMES, MESSAGE, OSeMOSYS, LEAP, PyPSA; TIMES-NZ, RETA
Hybrid	Integrates macroeconomic drivers with technology modules	Balances economic consistency with engineering realism; supports cross-sector trade-offs	Computationally demanding; transparency issue; inherits both paradigms' limits	MESSAGE-MACRO, TIMES-MACRO, REMIND; NZ soft-linking TIMES-NZ with CGE

Taken together, structural paradigms reveal that no single approach adequately captures the multi-scale realities of energy transitions. Top-down models provide breadth but lack engineering depth; bottom-up models offer granularity but often miss economy-wide interactions; and hybrid frameworks attempt to reconcile both, though at the expense of data intensity, computational load, and interpretability. These tensions are particularly evident when hybrid designs are applied to long-lived investment decisions, where policy, technology, and market uncertainties interact in ways that aggregated assumptions cannot easily resolve.

For industrial transitions such as process heat decarbonisation, hybrid frameworks are especially relevant because they link sector- and site-level choices (e.g., biomass or electrification options) with economy-wide drivers (e.g., ETS price volatility, macroeconomic feedbacks). Yet the same integrative qualities that make them powerful can also constrain their usability in practice, as complex coupling across scales often reduces transparency and hampers stakeholder confidence. Addressing these limitations requires methodological innovations that retain the integrative capacity of hybrid designs while enhancing tractability and openness. Emerging modular and open-source architectures exemplify this direction, offering ways to decompose complex systems into smaller, auditable units that can be recombined for multi-scale analysis [197].

Such advances point toward next-generation modelling paradigms that are not only structurally integrative but also explicitly oriented toward decision support under uncertainty. In this sense, structural classifications provide a foundation, but further insight requires attention to methodological orientations that illustrate how models are designed to describe, prescribe, explore, or stress-test transitions.

4.3 Methodological classification

While structural paradigms distinguish models by their scope and boundaries, methodological paradigms determine how systems are represented and solved, building on the epistemic foundations of uncertainty and decision-making [18]. They define the mathematical and conceptual philosophies that underpin model construction, shaping the nature of insights that can be generated, particularly for process heat decarbonisation, where site-level granularity and deep uncertainty demand robust, hybrid approaches [218]. Broadly, three methodological orientations dominate the energy modelling literature; optimisation, simulation, and equilibrium. These reflect contrasting views of energy systems, as decision problems (e.g., least-cost boiler retrofits at site), as evolving processes (e.g., regional biomass supply interaction), or as markets seeking balance (e.g., ETS-driven fuel equilibrium). Many contemporary tools blend elements of these paradigms, but their distinct logics remain useful for classification and critical assessment [93], informing the synthesis of positivist rigor (optimisation/equilibrium) with postpositivist exploration (simulation) in next-generation frameworks [20], [153].

4.3.1 Optimisation: energy as a decision problem

Optimisation has been the cornerstone of energy system modelling because it formalises the search for least-cost or least-emission pathways under resource, technical, and policy constraints [224].

Building on the positivist traditions discussed in Section 1.2, where models serve as objective tools for pathway identification, optimisation paradigms treat energy transitions as decision problems solvable through mathematical programming. The most widely used formulations are linear programming (LP), mixed-integer linear programming (MILP), and nonlinear programming (NLP), each reflecting different assumptions about convexity, continuity, and discreteness. Together, they define how decisions are represented and solved, from aggregated long-term transitions to site-level technology replacement in contexts like New Zealand's industrial process heat sector.

LP has traditionally underpinned capacity-expansion and long-term planning models, such as TIMES or MARKAL. Its convex and continuous structure guarantees tractable solutions at national and multi-regional scales, with the added benefit that dual variables (shadow prices) can be interpreted as marginal costs of system constraints [93]. LP is particularly well suited to multi-decadal transition analysis, such as exploring aggregate fuel switching or national carbon budgets under ETS caps [4]. However, it assumes divisibility and linearity, meaning technologies are represented as continuous flows. While acceptable for aggregate planning, this is less realistic for industrial process heat, where indivisible assets such as boilers, furnaces, or dryers must be captured explicitly.

To address indivisibilities, MILP extends LP by introducing binary and integer variables for discrete investment and operational choices. This makes MILP indispensable for representing industrial utilities and process equipment, such as the on/off operation of boilers, or staged retrofits of electric heat pumps. In the process heat context, MILP enables explicit modelling of technology replacement decisions, for example switching from coal boilers to electric alternatives in discrete unit sizes at sites like Edendale, where grid constraints (GXPs) and emissions limits must be enforced [16]. The drawback is computational; the inclusion of integer constraints increases complexity sharply as temporal or spatial resolution grows. To address this, decomposition or two-level optimisation strategies (continuous investment planning solved first, followed by detailed operational scheduling) are often used to maintain tractability [225].

When system dynamics cannot be adequately represented by linear approximations, NLP becomes necessary. Typical sources of nonlinearity include part-load boiler efficiency curves, thermodynamic coupling in CHP systems, and non-linear relationships in heat exchanger networks. NLP captures this engineering realism more faithfully than LP or MILP, but it comes with computational challenges and the risk of converging to local rather than global optima [164], [216]. In practice, many process heat applications linearise nonlinearities using piecewise-linear approximations, embedding them back into LP or MILP formulations. Advanced cases, such as modelling detailed CHP or industrial heat recovery

systems (e.g., in dairy processing), may require true nonlinear methods or hybrid approaches combining NLP with metaheuristic techniques.

Several methodological extensions have been developed to enhance realism and robustness. Multi-objective optimisation allows simultaneous consideration of cost, emissions, and other goals, using approaches such as weighted-sum functions or Pareto frontiers [226]. In industrial energy planning, this enables balancing decarbonisation targets with competitiveness, employment, or reliability (e.g., cost-emissions trade-offs for heat pump adoption under regional constraints).

Stochastic programming incorporates probability distributions for uncertain parameters, such as fuel prices, demand growth, or renewable output, producing expected-value or risk-adjusted solutions [134], [218]. Robust optimisation, by contrast, seeks solutions that remain feasible and near-optimal across a wide range of uncertain futures without requiring explicit probabilities [176]. This approach is particularly relevant to industrial process heat, where biomass availability, carbon pricing, and electricity tariffs are highly uncertain, and therefore required stress-testing against epistemic/deep uncertainties.

Large-scale optimisation models, especially those with high temporal and spatial resolution, are often computationally intractable without decomposition techniques. Approaches such as Benders decomposition or Lagrangian relaxation partition the original problem into coordinated sub-problems, each tackling a structural feature such as unit operations, time stages, or uncertainty realisations [226]. These methods are especially valuable when coupling process heat unit commitment with regional electricity balancing, where a single MILP may otherwise be unsolvable. Rolling-horizon and myopic formulations add further realism by relaxing the assumption of perfect foresight. Instead of solving for the full horizon, they resolve sequential short-term problems, updating decisions as new information becomes available [226]. This staged perspective mirrors how firms implement industrial retrofits under evolving policy and market conditions and provides a bridge between deterministic optimisation and robust decision-making, informing the hybrid approach required for decision-making in decarbonising process heat context.

4.3.2 Simulation: energy as a dynamic process

Simulation paradigms conceptualise energy systems not as optimisation problems seeking least-cost or least-emission outcomes, but as evolving processes shaped by dynamics, feedbacks, and interactions. They are particularly suited to contexts where uncertainty, heterogeneity, and non-linear behaviour dominate, and where the aim is to explore system behaviour rather than prescribe an optimal solution. Unlike optimisation, which treats energy as a decision problem under constraints,

simulation emphasises emergent properties and adaptive evolution. This aligns with postpositivist epistemologies in modelling (as discussed in Section 1.2), viewing models as tools for illuminating trade-offs under deep uncertainty rather than discovering singular truths [153]. In industrial process heat decarbonisation, simulation complements deterministic optimisation by capturing site-level variability (e.g., stochastic demand at dairy sites like Edendale) and regional feedbacks (e.g., biomass supply constraints in Southland).

System Dynamics (SD) represents energy systems as interconnected stocks, flows, and feedback loops, capturing how resources, technologies, and policies interact over time [185]. It has been extensively applied to energy and climate transitions since the 1970s, focusing on lock-in effects, rebound dynamics, and delayed policy impacts [184]. Its strength lies in conceptual clarity and transparency, which by highlighting reinforcing and balancing feedbacks, SD exposes unintended consequences of interventions. For New Zealand's process heat sector, SD can illustrate how investment cycles, ETS pricing volatility, and GIDI subsidies interact to generate sectoral inertia or path dependency in dairy manufacturing. However, abstraction is both its strength and limitation. While SD provides valuable insights into system-wide dynamics and policy timing, it lacks engineering detail and site-specific fidelity, making it most effective when integrated with bottom-up or optimisation frameworks [93], [104].

Agent-Based Modelling (ABM) simulates the behaviour of autonomous agents (households, firms, utilities, or regulators), whose individual decisions and interactions generate collective system outcomes [227]. Unlike optimisation or equilibrium models that assume representative agents and perfect foresight, ABM captures bounded rationality, behavioural diversity, and adaptive learning. Its role as a distinct simulation paradigm has been emphasised in recent reviews [93], [228], particularly for studying technology diffusion, market design, and heterogeneous responses to policy. For industrial process heat in New Zealand, ABM is particularly relevant for modelling how firms adopt technologies such as electric boilers or heat pumps under varying perceptions of risk, payback expectations, or RETA incentives, especially in heterogeneous sectors like dairy [16]. By capturing diffusion dynamics and potential tipping points (e.g., coal phase-out in Southland), ABM can reveal how collective behaviour shapes transition speed under epistemic uncertainty [18]. However, ABM faces challenges in empirical parameterisation (e.g., firm-level behavioural data) and computational burden and is often most powerful when combined with techno-economic optimisation in hybrid frameworks.

Discrete Event Simulation (DES), by contrast, is generally treated in the literature not as a separate paradigm but as a specialised simulation method for event-driven processes [229], [230]. DES models represent systems as sequences of discrete events, updating states only when events such as

equipment failures, process starts, or dispatch decisions occur. This makes DES highly effective for operational and industrial contexts where sequencing and downtime are critical [231]. In process heat applications, DES can simulate thermal load scheduling, downtime during retrofits, or activation of backup boilers during demand peaks at industrial sites (like Edendale dairy plant), incorporating aleatory uncertainty in hourly demand. DES provides valuable operational fidelity (modelling standby, ramp-up, and full-load states) but is less suited for long-term transitions and requires detailed process-level data, which is often scarce. Thus, DES is best positioned as a complement to optimisation and SD, rather than a stand-alone paradigm.

Multi-scale simulation links different levels of detail, from component-level dynamics to regional or national energy systems [232]. By coupling site-level processes (e.g., boiler performance, storage operation) with higher-level drivers (e.g., regional electricity adequacy, national biomass supply), it provides a holistic perspective on how local constraints scale into system-wide impacts [93]. For process heat, multi-scale frameworks reveal how hourly site demand profiles and technology adoption interact with electricity peaks or seasonal biomass competition. Compared to aggregated optimisation frameworks like TIMES or MESSAGE, these approaches capture operational dynamics more faithfully, though at the cost of higher computational demands and integration challenges.

Taken together, simulation approaches provide a family of paradigms and methods that capture feedbacks, dynamics, heterogeneity, and operational sequencing. In the literature, SD and ABM are widely recognised as distinct simulation paradigms, while DES is better understood as a technique within simulation, albeit one with growing relevance for industrial energy systems. For process heat decarbonisation, SD clarifies feedback-level dynamics, ABM reveals behavioural heterogeneity, DES offers operational fidelity, and multi-scale models bridge site-to-region interactions. Collectively, they extend the scope of analysis beyond deterministic optimisation, enriching the capacity to design robust and adaptive decarbonisation pathways under deep uncertainty, directly informing the modular, exploratory principles of the next-generation framework developed in this thesis.

4.3.3 Equilibrium: energy as a balancing system

Equilibrium approaches conceptualise energy systems not as optimisation problems with defined objectives or as dynamic processes unfolding through feedbacks, but as self-regulating markets that supply and demand balance through price formation. Rooted in Walrasian general equilibrium theory [233], equilibrium is mathematically defined as a fixed point at which excess demand functions vanish, ensuring all markets clear simultaneously. This epistemic foundation reflects positivist traditions: agents are assumed to be rational and utility/profit maximisers, endowed with perfect information,

and operating within convex production and consumption sets[104], [234]. Such assumptions enable internal consistency and tractable welfare analysis but abstract from heterogeneity, bounded rationality, and disequilibrium phenomena that characterise real-world industrial transitions. In energy modelling, two major variants dominate: computable general equilibrium (CGE) models, which capture multi-sector, economy-wide interactions, and partial equilibrium (PE) models, which focus on specific energy or resource markets.

CGE models extend Walrasian principles to empirical applications by embedding microeconomic optimisation into large-scale numerical simulations [211]. They capture how shocks, such as carbon pricing, ETS reforms, or fuel taxes, propagate across production sectors, households, government, and trade. Producers are typically modelled as profit-maximising and households as utility-maximising, with substitution possibilities captured by constant elasticity of substitution (CES) functions [233]. In New Zealand, CGE approaches underpin models such as the New Zealand Treasury Model (NZTM) and C-PLAN, which have been applied to assess fiscal reforms, ETS trajectories, and climate policy impacts across agriculture, forestry, and industrial energy use [205], [206]. For industrial process heat, CGE models can explore how rising carbon prices or biomass subsidies affect national resource allocation, but they reduce site-level retrofits (e.g., whether a dairy plant adopts a heat pump or biomass boiler) to stylised substitution elasticities between capital, labour, and energy. This limits their ability to represent indivisible investments, technology lock-in, or site-specific constraints such as grid connection capacities. As Kornai (1971) and subsequent critiques emphasise, equilibrium's reliance on representative agents and perfectly competitive markets obscures heterogeneity and path dependency [234], [235], [236]. From a DMDU perspective, CGE's equilibrium assumptions make it poorly suited for stress-testing under deep uncertainty, as the framework presumes stability and convergence rather than disequilibrium dynamics.

PE models narrow the scope to a single sector or market, most commonly energy or emissions. They rely on supply–demand curves, marginal abatement cost schedules, or resource allocation functions to determine equilibrium prices and quantities. By focusing on one domain, PE models are more tractable and transparent than CGE, and they can provide quick insights into the relative costs of competing abatement options [211]. In New Zealand, PE methods underpin EECA and MBIE's use of marginal abatement cost curves (MACCs) to identify cost-effective decarbonisation opportunities in process heat, biomass substitution, and electrification [237]. For example, MACC analysis highlights the relative cost of replacing coal with biomass in Southland compared to electrification elsewhere. However, PE's simplification comes at the expense of cross-sectoral realism; modelling biomass solely through a supply curve may neglect its interactions with forestry exports, regional land-use change, or competing industrial demand [44]. Moreover, by assuming rational actors and market clearing, PE

frameworks overlook firm-level heterogeneity, behavioural inertia, and structural change. For process heat, this means that while PE models are useful for benchmarking marginal abatement costs, they cannot capture adoption bottlenecks, sectoral clustering, or uncertainty ranges (e.g., ETS price volatility or ±25% biomass supply shocks [63], [238]).

The principal value of equilibrium approaches in this study lies not in predicting industrial transitions but in providing consistency signals, such as, marginal costs, scarcity rents, and shadow prices, that can inform other paradigms. In next-generation modular architectures such as Project BlueSky, equilibrium-inspired dual variables provide a lightweight mechanism for linking regional resource availability (e.g., biomass supply, grid adequacy) with site-level optimisation and behavioural simulation [20], [93]. Reframed in this way, equilibrium ceases to be a stand-alone predictive paradigm and instead becomes a diagnostic layer embedded within hybrid, site-to-region frameworks. This is particularly relevant for New Zealand's process heat transition, where ETS-driven price signals must be balanced against local grid constraints and biomass competition [14]. Integrating equilibrium logics with optimisation and simulation thus enables a pluralist, postpositivist approach; optimisation identifies least-cost retrofit pathways, simulation explores behavioural and dynamic uncertainties, and equilibrium ensures internal consistency across resource markets. Together, these complementary logics support robust decision-making under deep uncertainty, laying the groundwork for adaptive, least-regret strategies in industrial decarbonisation.

The main value of equilibrium approaches lies in providing consistency signals (marginal costs, scarcity rents, or carbon prices). This aligns with next-generation modular architectures such as Project BlueSky, where equilibrium-inspired dual variables provide a lightweight mechanism for linking regional resource availability with site-level decisions [20]. In this way, equilibrium ceases to be a prescriptive model of "perfect markets" and instead becomes a diagnostic tool integrated with optimisation and simulation, supporting robust decision-making under uncertainty in New Zealand's process heat decarbonisation (e.g., balancing ETS-driven prices with grid constraints) [14].

Optimisation, simulation, and equilibrium represent complementary logics for energy system analysis. Optimisation offers normative insights—what *should* happen given objectives and constraints. Simulation provides exploratory insights—what *could* happen given dynamic processes and behavioural rules. Equilibrium delivers descriptive insights—what *would* happen if markets cleared under given policies. Yet, taken in isolation, each paradigm provides only a partial lens on the energy system. Their comparative strengths and blind spots highlight the need for pluralistic approaches, particularly under conditions of uncertainty and heterogeneity. Table 4.3 distils these contrasts, outlining the core logics, strengths, limitations, and representative models.

Table 4-3 Comparison of methodological paradigms in energy system modelling

Paradigm	Core Logic	Strengths	Limitations	Representative Models
Optimisation	Identifies least-cost (or least-emission) solutions given constraints	Provides clear policy or investment benchmarks; rich technology detail	Assumes perfect foresight; sensitive to input data; limited behavioural realism	TIMES, MESSAGE, PyPSA (planning mode)
Equilibrium	Represents market interactions through supply–demand clearing	Captures price formation, trade, and macroeconomic feedbacks	Limited technology detail; assumes competitive markets and equilibrium	CGE models, GEM-E3
Simulation	Simulates system behaviour over time using rules, accounting, or feedback loops	Represents dynamics, feedbacks, and emergent behaviour; useful for exploratory analysis	No single optimal solution; calibration challenges; can be data-intensive	LEAP, System Dynamics, Agent-Based Models

Taken together, these paradigms show that no single methodological lens can capture the heterogeneous and uncertain realities of energy transitions. Optimisation provides tractable insights into cost–emissions trade-offs but struggles with uncertainty and behavioural realism; simulation illuminates feedbacks and emergent dynamics but lacks prescriptive clarity; equilibrium ensures systemic consistency but at the expense of technological resolution and heterogeneity. For industrial process heat transitions, particularly in the New Zealand context, these trade-offs are acute as site-level diversity, regional resource constraints, and volatile policy environments require pluralistic and often hybrid approaches that selectively combine the strengths of each paradigm [197], [239], [240].

Looking forward, next-generation models must move beyond rigid paradigm boundaries and integrate these logics in ways that balance tractability, dynamism, and systemic breadth. In this view, optimisation's normative clarity, simulation's attention to uncertainty, and equilibrium's macroeconomic feedbacks are not substitutes but complementary building blocks. Their integration is essential for credible analysis of industrial decarbonisation, where long-lived investments must remain resilient to shocks, feedbacks, and structural change. This recognition highlights that methodological paradigms, while central, do not fully determine a model's policy relevance; equal weight must be given to how models are oriented toward decision-support.

4.4 Decision orientation classification

Building on the methodological lenses, an additional way of classifying models concerns their orientation toward decision-support. Whereas methodological paradigms emphasise how systems are represented and solved mathematically, decision orientations highlight the purpose of modelling and the kinds of insights it seeks to generate. This reflects the shift in philosophy from positivist notions of models as “truth-discovery tools” to post-positivist perspectives where models serve as structured devices for pluralistic exploration, negotiation, and decision-making under uncertainty [9].

In practice, this means that energy models can serve descriptive functions, documenting historical patterns and system characteristics; prescriptive functions, identifying cost-optimal or policy-optimal pathways [51]; exploratory functions, stress-testing strategies across a wide range of uncertain futures [9]; or normative functions such as backcasting, which work backward from long-term societal goals [241]. More recent developments have introduced robust and adaptive orientations that explicitly address deep uncertainty by focusing on resilience and least-regret strategies [9], reflecting a growing emphasis in both academic and applied domains, including IRENA's World Energy Transitions Outlook 2024 with 2025 updates on infrastructure resilience [242], and applications in national infrastructure stress-testing [51].

Decision orientation matters because it determines not only the analytical methods employed, ranging from linear programming to exploratory ensembles and adaptive pathways, but also the policy relevance and uptake of modelling outcomes [243]. For New Zealand’s industrial process heat sector, for instance, descriptive approaches underpin demand baselines such as EECA’s RHDD [7], prescriptive optimisation supports cost–emissions trade-offs under the Emissions Trading Scheme (ETS), and robust/adaptive approaches are increasingly required to address uncertainties in biomass availability, electricity prices, and policy stringency. Orientations also shape the stakeholder engagement process, since models intended for exploration or robustness must be designed for transparency, iterative learning, and legitimacy across diverse decision-makers.

The following review examines these orientations in turn, outlining their methodological underpinnings, strengths and limitations, and implications for the design of next-generation modelling frameworks. This provides a foundation for later discussion of modelling requirements and the framework proposed in subsequent sections.

4.4.1 Descriptive orientation

A descriptive orientation in energy system modelling is primarily concerned with representing and documenting the system as it currently exists or has evolved over time. Rather than prescribing optimal pathways or exploring uncertain futures, descriptive models provide reference points and empirical baselines that characterise system structure, historical demand, and technology stock. In futures studies, such models correspond to *forecasting* or *baseline scenarios*, which extend present trends forward without imposing normative targets or substantial structural changes [241]. Their main function is to support understanding of the system's configuration and dynamics under existing policies, institutions, and behaviours, offering a common factual ground for more normative or exploratory exercises [244].

Methodologically, descriptive approaches rely heavily on accounting frameworks, simulation, and econometric or statistical methods calibrated to observed data. Common techniques include energy balance tables, technology stock models, regression-based demand forecasts, and calibrated simulation of supply–demand dynamics [140], [197]. These methods often prioritise transparency and data consistency, ensuring that stakeholders can clearly trace the underlying assumptions and empirical foundations. Examples in energy modelling include the International Energy Agency's World Energy Balances [3], the Long-range Energy Alternatives Planning system (LEAP) in baseline mode [130], and national inventory systems aligned with UNFCCC reporting requirements [245].

The strengths of descriptive approaches lie in their transparency, empirical grounding, and communicative value. They provide essential benchmarks for validating more complex models, act as starting points for scenario design, and ensure that debates are anchored in empirically defensible data [9]. Moreover, descriptive baselines are critical for stakeholder engagement, offering a common frame of reference that is less value-laden than normative scenarios [10].

However, descriptive orientations also exhibit significant limitations. By design, they risk reinforcing "business-as-usual" assumptions and underestimating the likelihood of disruptive changes in technology, policy, or behaviour [241]. They also tend to be static, often lacking mechanisms to capture dynamic adaptation or emergent socio-technical shifts. As van Notten et al. argue, descriptive baselines can serve important communicative functions but should not be mistaken for forecasts of the most likely future, as their methodological simplicity limits their capacity to anticipate discontinuities [244].

In the New Zealand context, descriptive modelling provides the foundations for understanding current patterns of industrial energy use. National energy statistics and sectoral demand datasets compiled by MBIE and EECA establish empirical baselines for process heat and other end-uses, supporting

benchmarking and model validation [7], [223]. While indispensable for establishing a factual reference point, such descriptive approaches are insufficient for long-term decarbonisation planning in an environment characterised by volatile fuel costs, evolving emissions policies, and uncertain technology uptake, highlighting the need for prescriptive orientations to guide targeted interventions, as explored next.

4.4.2 Prescriptive orientation

A prescriptive orientation in energy system modelling is concerned with guiding decision-makers toward what *should* be done to achieve specified objectives, commonly framed around minimising system cost, reducing emissions, or meeting policy targets. Rather than documenting the system as it is, prescriptive models adopt a normative stance, focusing on how the system ought to evolve under defined goals and constraints. They are closely aligned with what Börjeson et al. (2006) term *normative scenarios*, where futures are explicitly designed around a desired end-state such as achieving a renewable energy share or reaching net-zero emissions [241]. In this way, prescriptive models translate philosophical debates on the purpose of modelling and the methodological tools of optimisation into a structured decision-support role; outlining cost-optimal or policy-consistent pathways [244].

Methodologically, prescriptive approaches rely on optimisation-based frameworks. Prominent implementations include the MARKAL/TIMES family [128], [139], and MESSAGE [156], which have been widely applied in national and international energy planning. These models typically integrate technology-rich representations with policy and resource constraints to determine least-cost expansion or transition pathways. Their outputs are often framed as “optimal” solutions given assumptions, which makes them particularly attractive for policymaking where clarity and quantification of trade-offs are highly valued [246].

The strengths of prescriptive orientations lie in their ability to translate complex systems into tractable decision problems. By systematically combining technical options with resource, economic, and environmental constraints, they provide clear signals about relative cost-effectiveness and the potential role of different technologies. This clarity supports strategic planning, target-setting, and investment prioritisation, making prescriptive models central to national decarbonisation roadmaps and climate policy assessments.

However, this orientation also has important limitations. Prescriptive outcomes are highly sensitive to assumptions about technology costs, policy instruments, and demand trajectories. By presenting a singular “optimal” pathway, they risk conveying a false sense of precision in contexts where the future is deeply uncertain [93], [247]. As van Notten et al. (2003) note, normative scenarios inevitably embed

value judgements and assumptions about what futures are considered desirable [244]. This makes them politically salient and powerful in policy debates, but also contested, since stakeholders may challenge the assumptions or criteria used to define "optimality". Furthermore, because prescriptive approaches often prioritise cost-effectiveness, they can underrepresent social, institutional, or behavioural factors that influence real-world decision-making. Their outputs are therefore best understood not as predictions but as conditional insights, valid only under explicitly stated assumptions [246].

In practice, prescriptive orientations have shaped many influential energy and climate assessments, from EU decarbonisation roadmaps to integrated assessment modelling. In New Zealand, the TIMES-NZ model has been used to examine least-cost decarbonisation pathways, including the future of industrial process heat [178]. These applications underscore the value of prescriptive approaches for providing structured guidance, while also highlighting the need to complement them with exploratory and robust orientations capable of addressing uncertainty and incorporating diverse stakeholder perspectives.

4.4.3 Exploratory orientation

Moving beyond descriptive baselines and prescriptive pathways, an exploratory orientation is concerned with examining *what could happen* across a wide range of plausible futures. In contrast to prescriptive optimisation, which focuses on *what should happen* under defined objectives, exploratory approaches deliberately shift attention to *what could unfold* when multiple uncertainties interact. This corresponds to Amara's (1981) distinction of *possible futures* [248] and Börjeson et al.'s (2006) category of *explorative scenarios*, often operationalised through large ensembles and systematic stress-testing [241]. In philosophical terms, this orientation reflects the movement away from deterministic, positivist modelling toward post-positivist perspectives in which models are used as heuristics for pluralistic exploration rather than as predictors of a singular truth [9]. It also resonates with the tradition of simulation approaches that emphasise feedbacks and dynamic complexity, but extends them by explicitly spanning uncertainty space, including epistemic and deep uncertainties.

Methodologically, exploratory orientations rely on tools designed to capture wide variability in system drivers. Techniques include Monte Carlo simulation, Latin Hypercube sampling, exploratory modelling [174], and scenario discovery methods that apply data-mining and statistical learning to identify the conditions under which strategies succeed or fail [249], [250]. These methods shift the focus from generating a single "optimal" trajectory to mapping the range of plausible system responses, thereby helping decision-makers recognise resilience, fragility, and policy blind spots. Exploratory approaches

are particularly suited to contexts where multiple drivers, such as, technological, economic, environmental, and behavioural, interact in unpredictable ways.

The strengths of exploratory orientations lie in their ability to illuminate structural vulnerabilities and broaden decision perspectives [10], [250]. By spanning a large possibility space, they reveal risks that would remain invisible under prescriptive or descriptive approaches and encourage more flexible and adaptive policy design [9]. They also enhance transparency by engaging stakeholders in the evaluation of diverse futures rather than privileging a single outcome [249].

At the same time, exploratory approaches face important limitations. The production and analysis of large ensembles can be computationally demanding, particularly when combined with detailed, technology-rich system representations [93]. The interpretation of multidimensional results is also challenging, where decision-makers may find themselves overwhelmed by the volume and complexity of scenarios [10]. Their usefulness therefore depends on careful framing, visualisation, and participatory processes to translate broad scenario sets into actionable insights [249]. In addition, scenario ensembles remain shaped by modeller choices regarding system boundaries and uncertainties considered, which can introduce subtle biases [244]. Finally, the absence of a single recommended pathway may reduce policy uptake, as stakeholders and decision-makers sometimes prefer concise prescriptions over open-ended explorations [247].

Despite these challenges, exploratory orientations have become highly influential in both academic and applied energy contexts. Shell's corporate scenarios pioneered their strategic use [99], while ensemble-based assessments in integrated assessment modelling (IAMs) now form a standard part of global climate policy evaluation, as seen in recent IPCC AR6 explorations [251], [252]. Most importantly, exploratory approaches underpin modern decision-making frameworks, where scenario discovery and robustness analysis are used to design strategies that remain effective across divergent futures [9], [10]. For example, In New Zealand's process heat sector, exploratory ensembles could map scenarios for ETS price volatility and biomass supply risks, informing resilient strategies beyond traditional baselines [16], [29]. In this way, exploratory orientations serve as a bridge between prescriptive optimisation and the robust/adaptive approaches that explicitly embed resilience under deep uncertainty.

4.4.4 Backcasting orientation

Building on exploratory openness to multiple futures, a backcasting orientation adopts a normative stance similar to prescriptive optimisation, but with a crucial difference in logic: rather than starting from present conditions and projecting forward, it begins with a desired long-term target and works

backward to identify the actions, policies, and system changes required to achieve it. First articulated by Robinson (1990) [253], and further developed in sustainability studies [254], [255], backcasting is particularly relevant where societal goals are clearly defined but the pathways to reach them remain uncertain, such as achieving net-zero greenhouse gas emissions by 2050. In this sense, it complements prescriptive models by retaining a normative orientation, while sharing with exploratory approaches an openness to multiple feasible pathways connecting present and future states.

Methodologically, backcasting combines pathway analysis, participatory scenario building, and policy road mapping. The process typically starts with the articulation of long-term objectives, often derived from international agreements or national policy targets. From there, modelling tools are used to test the feasibility of different pathways, identify intermediate milestones, and highlight structural transformations required across energy, infrastructure, and behaviour. Techniques can range from optimisation models constrained by end-goals to simulation models testing feasibility, to participatory workshops where stakeholders co-develop transition pathways [253], [254], [255].

The strengths of backcasting lie in its ability to align near-term decisions with long-term sustainability visions. By focusing on normative goals such as decarbonisation or renewable integration, it provides a clear anchor for policymaking and stakeholder engagement, ensuring that short-term policies are not evaluated in isolation but in relation to strategic objectives [241]. It is also effective at fostering collaboration, since stakeholders are invited to debate not only the desirability of goals but also the plausibility of alternative pathways to achieve them [254].

At the same time, backcasting faces important limitations. By assuming a fixed end-state, it may underestimate the challenges of feasibility, institutional inertia, or unforeseen technological disruptions [255]. Unlike exploratory approaches, it does not fully span the breadth of uncertainties, and unlike robust/adaptive methods, it does not explicitly build in mechanisms for adjustment when assumptions fail. Backcasting scenarios can therefore risk oversimplification if they neglect deep uncertainty or treat pathways as linear sequences of milestones rather than adaptive trajectories [244].

Despite these challenges, backcasting has been widely applied in energy and climate planning, from European Commission long-term strategies [256] to IPCC mitigation pathways [257]. In New Zealand, backcasting logic underpins the Climate Change Commission's 2024 review of the 2050 emissions target, where present policy levers are evaluated against net-zero goals [258], including process heat transitions through initiatives like the RETA program [16]. More broadly, the orientation illustrates how normative goals can be translated into structured pathways, clarifying the minimum system changes and milestones that must occur for ambitious targets to remain credible. Yet its limitations in

addressing uncertainty and feasibility also highlight the need for approaches that embed adaptability and resilience alongside normative vision.

4.4.5 Robust and adaptive orientation

Robust and adaptive orientations extend beyond descriptive, prescriptive, exploratory, and backcasting roles by explicitly addressing conditions of *deep uncertainty*. These are contexts where models, probabilities, or values cannot be agreed upon [137]. Unlike prescriptive optimisation, which seeks optimal pathways under fixed assumptions, or exploratory ensembles that stress-test across multiple futures, robust and adaptive approaches emphasise strategies that remain effective across a wide range of plausible conditions, including shocks and surprises [259]. Their purpose is not to identify one "best" outcome but to illuminate pathways and decision rules that minimise vulnerability while preserving flexibility over time [243].

A defining characteristic of this orientation is its sequential logic. Energy transitions are conceived as chains of interdependent decisions in which early investments shape or constrain subsequent options. Robust and adaptive models therefore focus on adaptive pathways (staged strategies that evolve as new information emerges) ensuring that near-term choices keep multiple long-term futures open [259]. For instance, a dairy plant might initially invest in efficiency upgrades, preserving flexibility to adopt electric boilers if grid capacity expands or biomass if supply chains stabilise. This perspective contrasts with one-off optimisation or static scenario logics, aligning instead with the reality of socio-technical transitions, where lock-ins, feedbacks, and path dependencies strongly influence outcomes. Studies of developing energy systems reinforce this view, highlighting the importance of staged, path-dependent decision-making when deterministic forecasts lack credibility under structural and institutional uncertainty [193].

In practice, robust and adaptive orientations have gained traction in both climate and infrastructure planning. International assessments increasingly employ resilience metrics and stress-testing, such as IRENA's World Energy Transitions Outlook, which emphasises infrastructure resilience [36], while adaptive pathway frameworks are embedded in national strategies to accommodate uncertainty in technology diffusion, fuel supply, and policy ambition [9]. For industrial process heat in New Zealand, such orientations are particularly relevant. Volatility in biomass supply, electricity prices, and ETS trajectories means that investments in long-lived assets must be staged and revisited as conditions evolve. By framing decarbonisation as an adaptive, iterative process, robust and adaptive approaches provide a powerful paradigm for guiding industrial transitions under uncertainty [93].

Despite their strengths, robust and adaptive approaches face challenges. They are computationally intensive, requiring large scenario ensembles or multi-stage optimisation, which can strain resources, particularly for high-resolution models [10], [19]. Their credibility also hinges on stakeholder engagement and transparent communication of assumptions; without these, outputs risk being perceived as opaque or overly complex. Nevertheless, practical applications demonstrate their value, including RAND's Robust Decision Making (RDM) in defence and climate policy [260], [261]. Deltares' Dynamic Adaptive Policy Pathways (DAPP) in Dutch water management, and emerging energy-sector models integrating adaptive stress-testing into transition planning [247], [262].

To synthesise these decision orientations and situate robust and adaptive approaches within the broader energy modelling landscape, Table 4.4 compares their core logics, strengths, limitations, and representative applications.

Table 4-4 Comparison of decision orientations approaches in energy system modelling

Orientation	Core Logic	Strengths	Limitations	Representative Models/Applications
Descriptive	Represent current/historical system	Transparent; empirical baseline	Business-as-usual bias; static	IEA balances, LEAP baseline, EECA RHDD
Prescriptive	Define cost/policy-optimal pathways	Clear policy signals; trade-offs	Sensitive to assumptions; narrow	TIMES, MESSAGE, TIMES-NZ
Exploratory	Stress-test across uncertainty space	Reveals risks; widens futures	Computationally heavy; complex	Scenario discovery, IPCC AR6 ensembles
Backcasting	Work backwards from end-goals	Aligns short- and long-term; participatory	Fixed targets; limited adaptability	EU strategies, NZ 2050 pathways
Robust & Adaptive	Resilient, flexible strategies over time	Handles deep uncertainty; sequential logic	Data- and compute-intensive; needs transparency	RAND RDM, Deltares DAPP, emerging models

Together, decision orientations demonstrate that the value of a model lies not only in its mathematical formulation but also in the type of decision support it provides. Descriptive tools ground analysis in empirical baselines; prescriptive optimisation identifies conditional pathways; exploratory and backcasting approaches broaden the futures considered; and robust/adaptive orientations emphasise resilience and sequential flexibility. For industrial process heat, where long-lived investments must remain credible under volatile and uncertain conditions, no single orientation is sufficient. Instead,

combinations are required that balance clarity, exploration, and robustness, underscoring the need for next-generation frameworks that can integrate these orientations in a transparent and tractable way.

4.5 Additional classification dimensions

While structural, methodological, and decision-oriented classifications capture the dominant ways of categorising energy system models, a further set of extended dimensions strongly shapes how models are applied, interpreted, and taken up in practice. These dimensions are not paradigms in the strict sense but rather design features that influence the scope, credibility, and legitimacy of modelling exercises. They determine how models are embedded in decision processes, shaping both the technical boundaries of analysis and the trust placed in their outputs by policymakers, industry stakeholders, and the public.

4.5.1 Purpose and function

Purpose and function form one such dimension, distinguishing models designed for forecasting near-term demand, planning long-term capacity, optimising least-cost pathways, or supporting decision-making under *deep uncertainty*. Clarity of purpose is essential, as emphasised in scenario typology literature [241], [244]. Van Notten et al. (2003) stress that scenario classifications depend on project goals, process design, and content [244], while Börjeson et al. (2006) distinguish between models that ask *what will happen*, *what can happen*, *or how a target can be reached* [241]. Forecasting approaches extend present trends into the future [93], planning models map long-term infrastructure and investment needs [263], optimisation models prescribe least-cost pathways under specified constraints [215], and decision-support models stress-test strategies across uncertain futures [243]. Explicitly defining purpose is therefore essential for aligning expectations and avoiding misinterpretation of results and whether outputs are to be treated as predictions, conditional prescriptions, or exploratory heuristics. In the context of decarbonising process heat, forecasting models provide demand baselines [16], planning models map infrastructure needs for fuel switching or electrification, optimisation models identify least-cost technology mixes under ETS constraints [111], and decision-support models stress-test these pathways against uncertainties, for instance, decision-support models in RETA assess biomass vs. electrification pathways under ETS price volatility (±25% by 2030) [4], [16]. This purpose-driven focus informs the spatial and sectoral scope of models.

4.5.2 Spatial and sectoral resolution

The geographical and sectoral scope of an energy system model is a critical design dimension that conditions the insights produced. Global integrated assessment models (IAMs) prioritise breadth,

linking energy, economy, and climate systems, but often lack detailed process-level representation [93]. National and regional models provide finer granularity, though they may simplify industrial heterogeneity or neglect site-specific constraints [200]. Site-level models deliver high fidelity, essential for analysing process heat transitions, but risk losing systemic perspective unless embedded within broader frameworks. For industrial decarbonisation, spatial resolution is particularly salient, where outcomes depend not only on facility-level technology adoption but also on regional resource availability, infrastructure capacity, and national emissions policies. Sectoral scope interacts closely with geography, as process heat cannot be meaningfully assessed in isolation from electricity and biomass supply systems. Capturing substitution, competition, and cross-sectoral feedbacks therefore requires an integrated treatment of both dimensions, moving beyond narrow feasibility studies to credible system-wide transition pathways [264]. These spatial considerations also interact with temporal resolution, shaping how models capture dynamics over time.

4.5.3 Temporal resolution and horizon

Temporal resolution and horizon define the time scales at which models operate, conditioning their capacity to link short-term operational feasibility with long-term transition pathways. High-resolution models (e.g., sub-hourly or hourly) capture load variability, storage dynamics, and renewable intermittency with fidelity, while coarser time steps or aggregated slices may suffice for long-term investment and policy strategy. Pfenninger (2017) demonstrates that reducing temporal detail through clustering or typical days can improve computational tractability but risks distorting correlations between demand and renewable generation, particularly under high renewable penetrations [199]. Likewise, Collins et al. (2017) show that multi-model approaches, combining long-term system models with detailed dispatch models, correct for underestimation of integration challenges like curtailment, congestion, or system inertia [265]. For industrial process heat, temporal scales matter at multiple levels. Seasonal variation in fuel demand is critical for biomass supply planning, daily and weekly cycles shape boiler and heat pump operation, and hourly resolution is essential to capture the interaction with electricity markets. At the same time, process heat decarbonisation involves investment horizons spanning decades, meaning models must reconcile operational realism with long-term infrastructure planning. Striking this balance often requires hybrid strategies, where high-resolution demand and supply characterisation at the site or regional level, embedded within broader multi-decade models. In this sense, temporal resolution is not a neutral choice but a determinant of whether models can credibly inform both operational and strategic contexts. Temporal dynamics, in turn, influence market behaviour, which governs actor interactions.

4.5.4 Market behaviour

Market behaviour shapes how competition, strategic interaction, and decision heterogeneity are represented. Classical energy system models often assume perfect competition, where firms act as price-takers and market-clearing occurs through equilibrium conditions. with firms as price-takers and market-clearing through equilibrium conditions, simplifying optimisation and equilibrium formulations [195]. However, real-world dynamics (strategic bidding, risk preferences, asymmetric information) require more nuanced approaches [93]. In electricity markets, optimisation capture single-firm profit maximization, equilibrium models represent strategic interactions (e.g., Cournot or supply function equilibria), and simulation models explore adaptive behaviours [266]. Agent-based modelling (ABM) represents heterogeneous actors adapting over time, effectively exploring technology diffusion and market resilience [266], [267]. Recent ABM approaches have advanced to incorporate AI-driven behavioural analytics, improving predictions of technology adoption under policy uncertainty [268].

While process heat decarbonisation is less directly exposed to wholesale electricity markets than power-sector investment studies, behavioural representation remains relevant across all fuels. Fossil fuel use is shaped by volatile international commodity prices and ETS-driven carbon costs [4]; biomass availability depends on regional feedstock competition, logistics, and policy incentives [16], [269]; and electricity adoption is increasingly conditioned by time-of-use tariffs, peak pricing, and wholesale price variability [158]. Technology adoption, investment timing, and responses to ETS trajectories all depend on how firms perceive cost risks, competition, and regulatory credibility [259]. Simplified competitive assumptions may suffice for stylised least-cost optimisation, but more realistic behavioural lenses, such as agent-based adoption models or bounded-rational investment frameworks, become important when evaluating uptake under uncertainty [266]. In this way, market behaviour interacts with purpose, geography, and temporal resolution, conditioning whether models provide stylised efficiency insights or capture the diversity of real-world industrial decision-making. Beyond these dynamics, other factors such as data intensity and transparency would determine model credibility and uptake.

4.5.5 Data intensity, transparency, and openness

Energy system models vary in their level of technological detail and data requirements. Technology-rich, optimisation-based frameworks such as TIMES and MESSAGE provide process-level fidelity and support detailed sectoral analysis, but demand large datasets on technologies, costs, and constraints, making replication and validation challenging [156], [215]. By contrast, more aggregate accounting or simulation-oriented models such as LEAP can operate with modest data requirements, though often

at the expense of technical realism and sector-specific credibility [130], [194]. This trade-off is especially salient for industrial process heat, where detailed data on boiler stock, fuel use, and operating conditions are scarce yet critical for credible transition analysis.

These tensions have motivated the development of modular and hierarchical approaches, where the level of detail is scaled to the research question; for example, nesting site-level industrial models within regional or national optimisation frameworks [20]. At the same time, openness and transparency have become recognised as equally important dimensions of credibility. Open-source models such as OSeMOSYS demonstrate how making code and assumptions visible can reduce entry barriers, support reproducibility, and build trust [21], [270]. In contrast, proprietary or opaque models, while sometimes offering advanced capabilities, risk undermining legitimacy in contested transitions [247], such as industrial decarbonisation.

For process heat specifically, striking this balance is vital. Technology-rich detail enables robust evaluation of alternatives such as biomass boilers or electrification, while modular design ensures tractability. Equally, transparent databases, version control, and reproducible workflows are now seen as best practice, allowing models to act not only as analytical tools but also as shared platforms for stakeholder engagement and policy learning [200], [246].

4.5.6 Emerging approaches

Beyond established structural, methodological, and decision-oriented traditions, a new wave of emerging approaches is reshaping how energy system models are conceived and applied. These approaches leverage advances in artificial intelligence (AI), machine learning (ML), and deep learning (DL), alongside generative methods and digital representations (e.g., digital twins), to enhance flexibility, automate knowledge extraction, and expand the range of futures that can be explored. While not yet stabilised into formal paradigms, their growing prominence in the literature signals a gradual methodological shift, blurring the boundaries between optimisation, simulation, and exploratory frameworks while embedding richer data-driven capabilities [93], [197].

AI and ML are increasingly deployed to manage the complexity and scale of modern energy systems. By analysing large datasets from sensors, markets, and weather records, ML algorithms can uncover non-linear patterns invisible to traditional models, act as surrogates to approximate computationally intensive optimisation problems, and generate synthetic data where empirical coverage is sparse [271]. Applications already span renewable forecasting (e.g., solar and wind variability), short- and long-term demand prediction, anomaly detection, and resilience assessment across multi-energy systems [272]. Recent advancements include generative AI for renewable forecasting, improving load

prediction and storage management in variable systems [273]. DL architectures, such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, have proven especially effective for high-dimensional time-series tasks like building energy demand and renewable output forecasting [274]. For process heat, such methods can fill data gaps in demand profiling, improve scenario ensembles where empirical monitoring is limited, and support optimisation of electrification or biomass adoption strategies. While ML/DL approaches achieve high predictive accuracy, they often lack interpretability and causal transparency, raising concerns when applied to long-term planning and policy evaluation. Their reliance on large, high-quality datasets further limits applicability in contexts with scarce site-level data [271], [274]. To mitigate these challenges, hybrid approaches are gaining traction; embedding ML components within techno-economic optimisation or simulation frameworks combines predictive power with system-level interpretability and robustness [271], [272]. In this sense, AI/ML are not replacements for knowledge-driven models but complementary tools, enhancing forecasting and operational analysis while extending uncertainty-aware planning.

Another interesting domain is the development of Digital twins, which are dynamic, data-driven replicas of physical assets or systems that combine real-time data streams with physics-based modelling and ML. They are increasingly used in energy systems for continuous monitoring, predictive maintenance, and short-term optimisation. In the power sector, digital twins simulate wind farms, solar assets, or grid operations with high fidelity, supporting resilience planning and adaptive market testing [275]. In industrial contexts, including process heat, digital twins are being explored to evaluate technology retrofits, electrification strategies, and demand-side flexibility under operational uncertainty [272], [276]. Their strength lies in operational fidelity and adaptability, supported by IoT sensor networks and high-resolution data capture that allow fine-grained simulations of system dynamics [277]. However, digital twins face limitations in long-term planning. Their reliance on high-frequency, real-time datasets makes them computationally intensive and difficult to extrapolate credibly over decades [275]. While they excel at capturing short-term variability, they struggle to represent structural or policy-driven transitions such as infrastructure build-out, fuel switching, or long-lived investment cycles [277]. The complexity of integrating diverse data streams also raises risks of reduced transparency and stakeholder trust if not carefully managed [277]. As such, digital twins are best deployed in hybrid frameworks where near-term operational insights complement system-level models like TIMES, MESSAGE, or PyPSA, which abstract detail to assess long-term investment and policy pathways [93], [228]. In this way, digital twins provide valuable operational realism but must be integrated with scenario-based and optimisation approaches to ensure credibility for strategic decarbonisation planning.

Generative AI extends the capabilities of energy modelling by creating synthetic scenarios, datasets, and techno-economic pathways, thereby reducing reliance on sparse or costly empirical data. Unlike traditional forecasting or optimisation, generative methods learn high-dimensional distributions and can generate plausible weather patterns, demand trajectories, or grid configurations conditioned on limited inputs [272]. Recent applications include enhancing meteorological datasets via generative adversarial networks (GANs), producing high-resolution renewable forecasts, and generating building energy load profiles when historical data are scarce [273], [278]. Such approaches expand the uncertainty space explored, complementing ensemble-based exploratory frameworks and could supporting robust decision-making under deep uncertainty [279]. For energy transitions, their value lies in stress-testing pathways and designing adaptive strategies where empirical data are insufficient. In building and industrial contexts, generative models can synthesise demand profiles for under-monitored facilities, simulate rare but high-impact weather or price events, and accelerate exploration of technology adoption pathways [278]. However, these models demand large, high-quality datasets, often struggle with physical interpretability, and may generate outputs that lack engineering validity. To retain policy relevance, generative AI is best integrated into hybrid frameworks, coupling synthetic scenario generation with physics-based or optimisation models (e.g., TIMES, MESSAGE, PyPSA), so that operational flexibility is preserved while structural, long-term planning constraints are respected [279].

These emerging approaches signal a shift away from static and deterministic representations toward more dynamic, data-rich, and uncertainty-aware frameworks. They do not supplant established paradigms but extend them, offering hybrid capabilities that combine optimisation, behavioural simulation, and adaptive decision support. At the same time, their limitations, including high data demands, limited interpretability, and risks of over-complexity, position their role in industrial decarbonisation as exploratory rather than fully established. For process heat, they represent promising yet still experimental avenues for capturing heterogeneity, addressing uncertainty, and fostering stakeholder engagement, while underscoring the continued need for transparent techno-economic foundations [271], [275].

The additional classification dimensions underline that the legitimacy and usefulness of energy system models depend not only on their structural or methodological core but also on design features such as purpose, resolution, behavioural assumptions, data intensity, and openness. Emerging approaches broaden this landscape, extending the options available to modellers while also raising new challenges of integration and governance. Table 4.5 summarises these extended dimensions, showing how different design logics shape the credibility, scope, and uptake of modelling exercises in industrial process heat decarbonisation.

Table 4-5 Summary of extended classification dimensions

Dimension	Categories / Options	Core Focus
Purpose & Function	Forecasting; Planning; Optimisation; Decision Support	Defines intended application and alignment with policy needs
Geography & Sectoral Scope	Global; National; Regional; Site-level; Single-sector vs. Multi-sector	Sets spatial boundaries and sectoral integration
Temporal Resolution & Horizon	Sub-hourly; Hourly; Seasonal; Annual; Short- to long-term	Captures operational dynamics and investment planning
Market Behaviour	Perfect competition; Imperfect competition; Oligopoly; Agent-based	Represents actor heterogeneity, market power, and strategic interaction
Data Intensity & Transparency	Technology-rich vs. Aggregate; Open-source vs. Proprietary	Balances technical fidelity with tractability, reproducibility, and trust
Emerging Approaches	AI/ML; Digital Twins; Generative/Hybrid Methods	Expands capabilities through data-driven, adaptive, and uncertainty-aware frameworks

The review of paradigms and extended dimensions illustrates that energy system models are never neutral technical devices. Their value depends not only on analytical rigour but also on the epistemic and design choices embedded in structural form, methodological style, decision orientation, and the additional dimensions discussed here. These choices determine not just the types of insights generated but also their credibility and legitimacy in the policy sphere, where models increasingly operate as contested instruments of transition.

Yet, as comprehensive as this typology is, it primarily explains how models are constructed and classified, rather than how they engage with the deep uncertainties and decision contexts that ultimately shape their relevance. Forecast-based optimisation or equilibrium approaches perform well when risks are quantifiable and futures relatively stable, but they falter under conditions of deep uncertainty, where disruptive technologies, volatile markets, and shifting policy landscapes dominate. Addressing these conditions requires moving beyond classification to frameworks that explicitly engage with uncertainty and decision support.

Chapter 5
Decision-Making under Deep Uncertainty: Foundations for a Decision-Centred Framework

5.1 Introduction

Energy system models are not merely analytical tools but decision-support instruments that operate within uncertain, dynamic, and often contested environments. Building on the classified modelling paradigms and orientations presented earlier, this section moves from taxonomy to the practical treatment of uncertainty for credible decision-making in industrial process heat decarbonisation.

A recurring challenge in decision science is the absence of a precise and universally accepted definition of a "decision". In optimisation, decisions are frequently represented abstractly as variables (e.g., investment x_t, operation u_t), with little interrogation of the construct [280]. Behavioural research focuses on how individuals evaluate alternatives under uncertainty rather than on formal definition [281]. Decision analysis has been more explicit. Howard (1966) defined a decision as an irrevocable allocation of resources [282]; a view well suited to capital planning but restrictive where choices are revisable (e.g., routeing under uncertainty) or concern methodological and informational commitments (e.g., deterministic versus probabilistic forecasting). Powell (2022) proposes a broader view: formally, a decision is an endogenously controllable information class; informally, "something we control". This re-centres controllability, rather than resource allocation, as the defining feature, aligning well with optimisation under uncertainty [280].

In modular energy models, site-level and regional choices therefore encompass not only investments and dispatch but also methodological and informational decisions (such as scenario-ensemble design, policy-lever selection, and data-provenance standards) that subsequently translate into physical implementation.

Classical decision theory frames a well-posed decision as comprising a set of alternatives, uncertain states of the world, consequences on valued attributes, and a preference structure. This leads normatively to expected-utility maximisation subject to constraints [283]. In practice, however, many real-world problems are ill-structured: goals are plural and contested, criteria are difficult to monetise, information is incomplete or ambiguous, and multiple solutions may be defensible. Jonassen (1997) therefore argues that such problems require justification rather than derivation, with success judged

by the transparent generation of alternatives, explicit criteria and weights, reasoned trade-offs, and a defensible choice within context [284].

Industrial process heat planning exemplifies these characteristics. Decisions involve long-lived, capital-intensive assets under uncertain biomass supply and delivered costs, GXP hosting constraints and reinforcement lead times, carbon-price trajectories, technology-cost learning, and organisational adoption dynamics. Governance introduces further uncertainty. In liberal, market-oriented systems such as New Zealand's, climate and energy policy combines market instruments (e.g., the Emissions Trading Scheme, ETS) with targeted interventions (e.g., the GIDI Fund and regional initiatives such as RETA). Electoral cycles, lobbying, and shifting priorities affect policy credibility, incentive strength, and implementation pace. These factors amplify deep uncertainty (disagreement about models, parameters, and values) and increase the premium on transparency, reproducibility, and stakeholder legitimacy [9], [137].

Credible analysis in this domain must therefore couple deterministic optimisation with exploratory evaluation within a unified workflow. Deterministic models remain indispensable for establishing physical feasibility and cost or emissions accounting with engineering fidelity, especially where electricity-system interactions are decisive, such as temporal chronology, GXP capacity, marginal emissions, and price signals. Temporal compression (e.g., representative-day selection) and multi-site aggregation (site → GXP/region → national) shape outcomes and must be transparently handled. These deterministic modules provide the dispatch, sizing, and constraint logs that anchor all recommendations.

Yet many decision-critical drivers in process heat lack reliable probability laws. This motivates an "explore-then-choose" stance that stress-tests candidate strategies across ensembles of plausible futures. Rather than pursuing narrow optimality, the focus shifts to robustness; supporting strategies that (i) satisfy performance thresholds across many futures (satisficing), (ii) minimise regret, (iii) limit downside risk, and (iv) expose conditions of vulnerability.

To translate analysis into actionable insight, the decision layer (including sensitivity analysis and multi-criteria evaluation) should be embedded rather than appended. Standard artefacts of each model run should therefore include robustness summaries (satisficing fractions, regret percentiles, reliability shares), weight- or assumption-sensitivity analyses, and MCDA-ready performance matrices, together with manifest files and acceptance checks to ensure auditability and provenance. An integrated, modular platform should interoperate with established power-system and resource modules where appropriate, while keeping interfaces transparent and well documented. Deterministic and

exploratory components should share data, constraints, and diagnostics through open, reproducible protocols—avoiding opaque or black-box coupling.

Finally, governance realities emphasise the importance of adaptive implementation. Rather than a one-off decision, effective planning requires pathways that employ signposts and triggers, such as ETS price bands, biomass-supply corridors, or GXP-upgrade milestones, to enable staged, revisable actions. Accordingly, this section establishes the uncertainty foundations and modelling implications; contrasts probabilistic tools with deep-uncertainty methods; introduces frameworks such as RDM, DAPP, and MORDM with their evaluation lenses; operationalises adaptivity through decision points, signposts, and triggers; and distils the implications for process heat decisions in the New Zealand context.

5.2 Types and sources of uncertainty

A central challenge in industrial process heat decarbonisation is that many of the most consequential factors for investment and policy decisions cannot be described by single "best-estimate" values. Following the classical distinction by Knight (1921), *risk* denotes situations where the distribution of possible outcomes can be credibly quantified [125], such as seasonal demand variability or boiler outage probabilities. By contrast, *uncertainty* arises where outcomes are recognisable, but probability distributions are unknown or contested [135], as in the case of biomass delivered costs or the learning rates of electrification technologies. At the extreme, *deep uncertainty* describes circumstances where stakeholders disagree on system models, parameters, or even on underlying values [9], such as the long-term credibility of the ETS or the pace of infrastructure reinforcement.

5.2.1 Types of uncertainty

The classification of uncertainty has long been a source of both intellectual fascination and practical debate. An anecdote recorded by E. T. Bell recounts Bertrand Russell's 1929 observation: "Probability is the most important concept in modern science, especially as nobody has the slightest notion of what it means" [285]. This captures a central difficulty. Although probability theory has become the dominant language for quantifying uncertainty, its interpretation remains contested [286]. In engineering contexts, an additional distinction must be drawn between uncertainties arising from inherent randomness and those due to incomplete knowledge [287]. he literature therefore commonly distinguishes aleatory, epistemic, and deep uncertainty as conceptually distinct categories [18], [179], [288]. While this typology cannot resolve all ambiguities, since the same variability may be classified

differently depending on the modelling frame, it provides a useful foundation for systematic treatment in decision-support contexts [10], [137].

- **Aleatory uncertainty** refers to inherent variability or randomness in a system, often regarded as irreducible. Its origins lie in stochastic phenomena such as dice rolls, the Latin *alea* means “chance”, and in engineering contexts it includes weather-driven variability, seasonal demand cycles, or the probability of a component failure. Aleatory uncertainty is best described through probability distributions and is particularly relevant for operational risk assessment and reliability analysis [292] [287].
- **Epistemic uncertainty** derives from limited knowledge or data gaps. It reflects what is not yet known, but could in principle be reduced by further measurement, research, or model refinement [288]. Examples include uncertain technology performance at scale, incomplete process-integration data at specific sites, or limited information on biomass logistics and regional yield factors [289]. Unlike aleatory variability, epistemic uncertainty can introduce dependence across system components if left untreated, for instance, correlated errors in parameter estimate that propagate across multiple facilities [287].
- **Deep uncertainty** arises when stakeholders disagree on models, probability distributions, or even on the objectives and values that should guide decisions [18]. This is common in long-horizon socio-technical transitions, where political, behavioural, and institutional dynamics resist probabilistic specification [9], [262]. In the process heat context, deep uncertainty surrounds the credibility of the New Zealand Emissions Trading Scheme (ETS) price trajectory, the long-term availability and cost of sustainable biomass under competing demands, and the timing of grid reinforcement at constrained GXPs [238]. Here the challenge is not measurement error but the absence of shared agreement on governing assumptions, which motivates decision frameworks focused on robustness and adaptability rather than point optimality.

5.2.2 Sources of uncertainty

Whereas the typology above clarifies the nature of what is unknown, an equally important perspective concerns where uncertainty enters the modelling process. The literature commonly distinguishes parametric, structural, contextual, and scenario uncertainties [179], [290]. This perspective identifies points in the modelling chain most vulnerable to variability, lack of knowledge, or disagreement, and clarifies which uncertainties may be reduced (e.g., through improved data) and which must instead be explored or stress-tested (e.g., policy volatility or divergent socio-economic pathways) [291].

- **Parametric uncertainty** refers to imprecision in the numerical values feeding into models, including technology capital and operating costs (CAPEX/OPEX), process efficiencies, demand forecasts, emissions factors, and delivered fuel prices [179], [291]. In process heat this appears as fluctuating biomass prices, uncertain learning rates for electrification technologies, and variability in site-level demand estimates.

- **Structural uncertainty** arises from how systems are represented within models. It stems from simplifications and modelling choices that shape feasible solutions [247], [291], including temporal aggregation (full chronology vs representative days), load-profile construction, site–region coupling, inclusion/exclusion of system constraints. These choices determine whether seasonal peaks, process integration, and regional bottlenecks are credibly represented.

- **Contextual uncertainty** concerns boundary conditions external to the optimiser but critical in practice, including policy rules, regulatory credibility, resource logistics, and infrastructure constraints [136], [291]. Examples include ETS policy design, biomass-logistics corridors, GXP hosting-capacity and reinforcement timelines. While sometimes modelled as fixed exogenous inputs, in reality these conditions evolve, making them important drivers of strategic risk.

- **Scenario uncertainty** is associated with alternative long-term futures, such as, socio-economic development, industrial activity, technology learning trajectories, and policy regimes. Unlike parametric or structural uncertainty, scenario uncertainty is not reducible to input ranges and requires narrative-based ensembles that test the resilience of strategies across fundamentally different pathways [9], [49], [262]. In the process heat context, this includes divergent futures such as high versus low export demand (e.g., in the dairy sector), rapid versus slow technology learning (cost reductions) for high-temperature heat pumps or contrasting carbon-price and fossil phase-out trajectories. These futures profoundly shape the scale and timing of technology adoption, investment risk, and regional energy-system integration.

Bringing these perspectives together clarifies method selection. Aleatory risk favours probabilistic or stochastic tools for near-term operations; epistemic and structural/contextual uncertainties motivate exploratory stress-testing with transparent structural toggles; deep uncertainty calls for robustness and adaptive pathways frameworks. In the New Zealand process heat context, these uncertainties often interact (for instance, policy credibility × grid-upgrade timing × biomass logistics), so model design must make their treatment explicit and traceable. A summary table (Table 5.1) can collate types and sources with concrete NZ examples to guide later modelling choices.

Table 5-1 Uncertainty Types and Sources in Process Heat Decarbonisation Modelling

Dimension	Definition	Process Heat Example
Types of Uncertainty		
Aleatory	Inherent variability, irreducible randomness	Hourly/seasonal demand fluctuations; random boiler outages; weather-driven biomass quality changes
Epistemic	Lack of knowledge, reducible with better data	Limited site-level metering; uncertain heat pump efficiency; incomplete biomass logistics data
Deep	Disagreement on models, distributions, or values	Divergent ETS price projections; contested biomass sustainability pathways; uncertain grid reinforcement timing
Sources of Uncertainty		
Parametric	Imprecision in numerical inputs	Fuel prices (biomass, electricity, coal, gas); technology CAPEX/OPEX; emissions factors; demand forecasts
Structural	Modelling choices and representations	Temporal aggregation (full chronology vs. representative days); synthetic load-shape construction; site–region coupling assumptions
Contextual	Exogenous boundary conditions and constraints	ETS policy design; biomass logistics radius and storage capacity; GXP hosting and reinforcement timelines
Scenario	Alternative futures beyond probabilistic specification	High/low biomass availability; rapid vs. slow grid reinforcement; stable vs. volatile carbon pricing regimes

The typology above underscores that process heat decarbonisation cannot rely on a single lens: least-cost optimisation is necessary but not sufficient. Credible decision-support must combine:

(i) deterministic optimisation to establish feasibility, binding constraints, and cost/emissions accounting;

(ii) probabilistic or stochastic methods where credible distributions exist to characterise operational risk;

(iii) exploratory, robustness-oriented approaches for deep and structural uncertainties; and

(iv) decision-evaluation layers such as MCDA to reveal trade-offs and stakeholder value judgements.

5.3 Deterministic and optimisation-based approaches

Deterministic optimisation has long underpinned energy-system analysis. In its canonical form, it identifies least-cost (or least-emissions) portfolios and dispatch subject to hard physical, policy, and

resource constraints [292]. Linear and mixed-integer formulations are particularly effective for capacity expansion, dispatch, and unit-commitment-style scheduling [224]. In the process heat domain, where investments are capital-intensive, site-specific, and tightly coupled to regional electricity and biomass systems, this layer is indispensable for:

- establishing feasibility with engineering fidelity;
- producing the artefacts that anchor downstream deliberation; and
- applying disciplined feasibility gates before any comparative ranking.

At the site scale, deterministic optimisation enforces heat and mass balances, unit operating limits (turn-down ratios, ramp rates, temperature lifts for high-temperature heat pumps), maintenance cycles, and integration constraints. At the system interface, it upholds binding external conditions that strongly shape technology adoption in New Zealand, including Grid Exit Point (GXP) hosting capacities and upgrade lead times; biomass logistics, storage, and sustainability constraints; and emissions accounting under ETS rules, including budget or band caps [16]. Because many operational risks materialise through chronology (e.g., seasonal peaks, storage cycling, or outage cover), temporal representation must be explicit. While chronology compression (e.g., representative days or weeks) can screen options, boundary or contentious cases should be re-evaluated at full-year resolution, accompanied by an auditable representativeness or error note [20].

Deterministic runs explain why outcomes emerge. They yield feasible or infeasible schedules, binding-constraint logs and shadow values (revealing scarcity, crowding, or congestion), and transparent cost and emissions ledgers. Aggregating results from site to region and national level preserves physical and policy interactions (who uses scarce grid headroom, where curtailment or storage binds, and how policy caps constrain portfolios), avoiding false confidence of isolated, site-by-site appraisals.

National and sectoral Energy System Optimisation Models (ESOMs), such as TIMES/MARKAL, MESSAGEix, and OSeMOSYS, and power-system planning tools (e.g., PyPSA, GenX) provide broader equilibrium and expansion context including carbon-price bands, technology cost bands, national biomass envelopes, indicative wholesale price and marginal-emissions trajectories, adequacy or reliability constraints [293]. The site to region layer returns bottom-up feasibility signals, such as, GXP hosting constraints and commissioning lead times, sustainable delivered-biomass within a logistics radius, and shadow-price diagnostics for scarce resources. Where mutual dependence is material, a small number of coordination passes (Gauss–Seidel on shared signals such as prices, headroom, and regional supply) can be used to align levels without constructing a monolith [20]. This keeps local

engineering detail where it matters, while maintaining national scenario consistency when policy or system conditions materially shift local signals.

Iterative coupling and gauss–seidel coordination

Large-scale energy systems are often too complex for monolithic optimisation, particularly when they integrate heterogeneous sub-systems with different time scales, data granularities, or model structures [294]. A widely used strategy is iterative coupling, in which sub-problems (e.g., site-level operations and regional supply constraints) are solved sequentially and exchange information until convergence. Among these, the Gauss–Seidel iteration provides a simple and interpretable coordination scheme that supports modular, multi-scale frameworks [20].

In cases where coordination is most naturally expressed through prices rather than quantities (e.g., a regional electricity module computes nodal prices as duals of balance/adequacy constraints and site models respond to those prices) modules can exchange dual variables iteratively. For α denoting to the dual(s) produced by the regional problem (e.g., price on a market-clearing constraint) and α_0 the price parameter fed into the site problem, the equilibrium of the two-module system coincides with the KKT (Karush-Kuhn-Tucker) conditions of a combined problem when $\alpha = \alpha_0$. Hence, an iterative scheme that updates α_0 with the newly computed $\hat{\alpha}$ and repeats until $\|\hat{\alpha} - \alpha_0\| \leq \varepsilon$ delivers a KKT-consistent fixed point (price consistency). In practice we use relaxation

$$\alpha_0{}^{(n+1)} \leftarrow (1-\lambda)\alpha_0{}^{(n)} + \lambda\, \hat{\alpha}^{(n+1)}, \lambda \in (0,1] \ ,$$

with residual and oscillation checks; when needed, Anderson acceleration or an augmented-Lagrangian/ADMM variant improves robustness [295]. This dual-passing variant complements the quantity-based Gauss–Seidel update and is particularly natural when regional modules expose reliable shadow prices for shared constraints.

Conceptually, suppose an energy system is decomposed into interdependent modules indexed by $m = 1, \dots, M$, each with decision vector x_m and shared coupling variables y (such as regional prices, capacities, or emissions). The general problem:

$$\min_{x_1,\dots,x_m} f(x_1, \dots, x_m)\, s.t.\, g_m(x_m, y) \leq 0, \forall m$$

can be solved by iteratively updating each module's solution while holding the others fixed. At iteration n:

$$x_m^{(n+1)} = arg\ \min_{x_m} f_m\left(x_m, y^{(n)}\right)$$

and the coupling variables are then updated according to an aggregation or reconciliation rule:

$$y^{(n+1)} = \Phi\left(x_1^{(n+1)}, \ldots, x_M^{(n+1)}\right)$$

where $\Phi(\cdot)$ represents consistency conditions such as total demand–supply balance, shared emissions budgets, or price formation. The process repeats until convergence of $y^{(n)}$ within tolerance ε.

The Gauss–Seidel procedure thus seeks a fixed point where each sub-problem's optimal solution remains optimal given the updated shared variables:

$$\left\|y^{(n+1)} - y^{(n)}\right\| \leq \varepsilon$$

This approach preserves decomposition transparency, enables parallelisation, and provides natural diagnostic points (residuals, dual variables, marginal costs) for analysing system-level feedbacks. It has been applied extensively in multi-region energy planning, electricity-heat coupling, and hybrid energy-economic models where full integration would be computationally intractable [296]. In the process heat context, such iterative coupling principles allow site-level deterministic optimisation to remain physically detailed while ensuring that regional supply modules (e.g., electricity and biomass) impose shared constraints through updated signals. This conceptual foundation prepares the conceptual ground for Chapter 3, which develops a modular framework that applies these principles to achieve consistent, scalable, and transparent coupling between local and regional model layers.

Credible deterministic analyses also incorporate acceptance checks within each iteration or run. Technical feasibility must be satisfied; reliability should remain within tolerance; emissions and budget constraints must hold; and solver convergence must be clean. When these conditions fail, diagnostics are recorded so that infeasibilities become sources of insight rather than silent attrition of runs [158].

Despite their rigour, deterministic approaches also have limits. Point inputs can render results brittle, and hidden structural choices, such as chronology compression, site–region coupling assumptions, the scope of electricity-system linkage, or static versus pathway formulations, can alter conclusions. The remedy is structural transparency: exposing such design choices as toggles, documenting them, and testing whether conclusions remain invariant across credible settings [291]. Deterministic outputs should therefore not be interpreted as forecasts but as feasibility and accounting baselines that anchor all subsequent uncertainty analyses [297].

For process heat planning in New Zealand, a fit-for-purpose deterministic layer should therefore:

- enforces physical and operational realities at site level;

- propagate regional constraints for electricity (time-series prices, marginal emissions, GXP headroom and upgrade steps) and biomass (availability by source and radius, seasonality, storage, and sustainability criteria);
- compute auditable key performance indicators (KPIs) for cost, emissions, peaks, and reliability; and
- emit complete provenance artefacts, configuration manifests, versioned signal and constraint packs, and acceptance flags, so that figures and tables can be traced to exact inputs, data sources, and code states.

Figure 5.1 summarises the decision-centred deterministic structure and data flows within the decision framework. Exogenous inputs enter first, including demand and load profiles (with synthetic profile generation where metered data are unavailable), ETS bands and policy caps, biomass logistics and sustainability screens, GXP headroom and upgrade steps, and time-series prices and marginal-emissions signals. These feed the site-level optimisation, which exchanges constraints and signals with regional modules (electricity and biomass). When available or required, a national model (e.g., TIMES-NZ or a PyPSA-based whole-of-grid study) supplies boundary conditions to the coordinator. Coordination is implemented via a Gauss–Seidel loop that updates shared signals and iterates with the site solver until a small tolerance is met; in practice this can be done in primal (quantity) space or dual (price) space by passing shadow prices from the regional module back to sites with simple relaxation. After convergence, acceptance checks gate the deterministic outputs (feasible schedules, KPIs, constraint or dual logs, and provenance artefacts), which then hand off to the uncertainty layers.

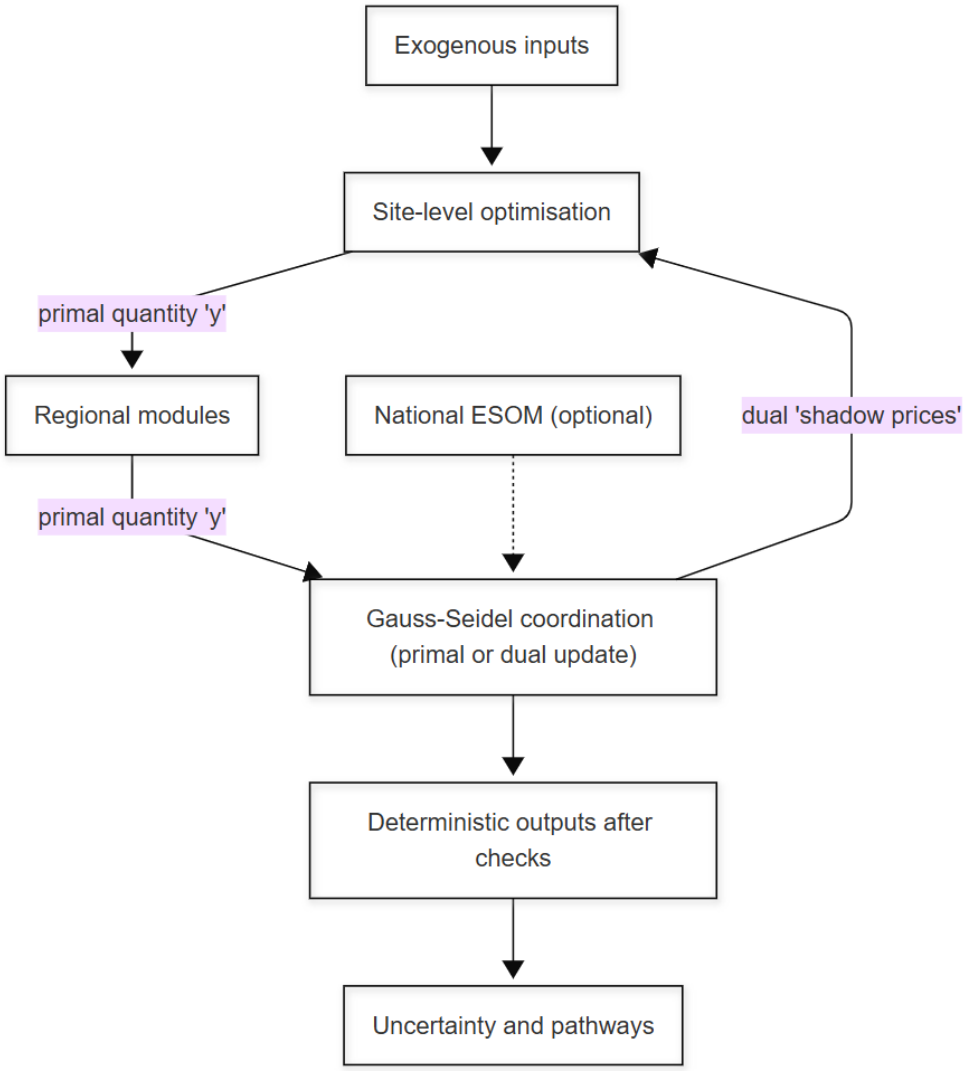


Figure 5.1 Role of the deterministic core within the decision framework

Beyond classical deterministic optimisation, stochastic programming provides a structured way to incorporate parametric uncertainty when probability distributions are credible [162]. Two-stage and multi-stage formulations separate *here-and-now* investment decisions from recourse actions, minimising expected or risk-weighted costs across scenarios [298]. Where uncertainty sets are bounded but not probabilistic, robust optimisation seeks decisions that remain feasible for all realisations within those bounds, complementing the deterministic core of process heat models [136]. With this foundation, subsequent layers (probabilistic, stochastic, and deep-uncertainty frameworks) can be coherently integrated on top of a transparent deterministic core.

5.4 Probabilistic and stochastic decision approaches

Probabilistic and stochastic approaches are widely used to represent uncertainty in engineering and energy-system models. They proceed on the premise that credible probability distributions (estimated

from historical data, high-quality forecasts, or expert elicitation) can be assigned to uncertain inputs, enabling statistical characterisation of outcomes and risk [299], [300]. In the process heat decarbonisation context, this premise is typically defensible for operational and near-term phenomena where measurement series exist or can be constructed with confidence. Examples include:

- **Fuel-price dynamics**, modelled with statistical time-series (e.g., autoregressive models) or continuous-time processes (e.g., Brownian or geometric Brownian motion [301]) to capture persistence and cross-price correlation [302];
- **Demand variability**, represented by probability distributions fitted to historical hourly/seasonal process heat loads (with allowance for heteroskedasticity) [303];
- **Technology performance**, by assigning distributions to parameters such as heat-pump COP, biomass-boiler efficiency, or availability [288]; and
- **Dependence structure** across inputs (e.g., fuel prices or across multi-site demand, using correlation matrices or copulas to capture tail dependence), so that portfolio effects are not mis-estimated [304], [305].

Where such probability distributions are credible, uncertainty can be represented with probabilistic simulation or stochastic-optimisation methods. Monte Carlo sampling (often enhanced through Latin Hypercube or quasi-Monte Carlo techniques for improved space-filling) propagates input uncertainty to model outputs by estimating expectations over sampled futures $\omega \in \Omega$ [306], [307]:

$$\mathbb{E}_{\omega}[.] \approx \frac{1}{N}\sum_{i=1}^{N}(.)(\omega_i)$$

When probability distributions are credible, for instance, for weather-driven resources or short-term market prices, stochastic programming provides a structured bridge between deterministic optimisation and robustness-oriented frameworks.

Two-stage and multi-stage formulations distinguish between here-and-now decisions (taken before uncertainty is realised, such as capacity or technology selection) and recourse decisions (adaptive operational responses after uncertainty unfolds). The general two-stage formulation is:

$$\min_{x} c^{\top}x + \mathbb{E}_{\omega}[Q(x,\omega)], \quad Q(x,\omega) = \min_{y_{\omega}}\{\, q_{\omega}^{\top}y_{\omega} : Wy_{\omega} = h_{\omega} - T_{\omega}x,\ y_{\omega} \geq 0\}$$

where x represents the investment or planning vector, y_{ω} the recourse actions under future ω, and T_{ω}, W define coupling constraints [134]. Multi-stage variants extend this logic across time, embedding successive recourse layers linked by state variables. In decomposition, the L-shaped (Benders) method exchanges duals from scenario subproblems to the master (as cuts), which is the stochastic analogue

of the dual-passing coordination; both enforce consistency with the KKT system of the underlying monolithic formulation while preserving modularity [295].

Chance-constrained formulations control the probability of constraint violation:

$$\Pr_{\omega}\{g_k(x,\omega) \leq 0\} \geq 1 - \alpha \quad ,$$

where x is the decision vector, $g_k(x,\omega)$ is the k-th constraint residual under future ω, and $\alpha \in (0,1)$ is a risk tolerance (e.g., $\alpha = 0.05$) [308]. Downside outcomes are controlled via Conditional Value-at-Risk (CVaR) at level α applied to a loss (or cost) variable $Z\ (x,\omega)$:

$$\mathrm{CVaR}_{\alpha}(Z) = \min_{\eta}\left\{\eta + \frac{1}{1-\alpha}\,\mathbb{E}_{\omega}[(Z(x,\omega) - \eta)^{+}]\right\} \quad , \qquad (u)^{+} = \max(u, 0)$$

With expectations typically evaluated through sample-average approximation over the ensemble of sampled futures ω_i [309].

These formulations are especially suited to aleatory, short-horizon risks, producing explicit risk metrics (e.g., downside quantiles, CVaR) that complement deterministic point results. Within an integrated framework, stochastic optimisation can also be implemented selectively, such as stochastic dispatch or storage scheduling at the site level, while keeping upper-level modules deterministic for tractability. In this sense, it serves as a transitional bridge, retaining the interpretability of deterministic optimisation while introducing controlled uncertainty to prepare for exploratory ensemble methods such as Robust Decision Making (RDM).

Figure 5.2 summarises the role of probabilistic and stochastic methods as an inner loop in the overall framework. When credible distributions are available, inputs feed two parallel branches: (i) Monte Carlo propagation and (ii) stochastic programming, whose L-shaped decomposition iterates between a master (first-stage) problem and scenario subproblems, passing primal first-stage decisions x to scenarios and returning dual information (cuts) until convergence. Both branches produce decision-relevant risk metrics (e.g., quantiles, CVaR, violation probabilities) that are (a) fed back to the deterministic core for feasibility/KPI reporting and (b) passed forward to RDM/MORDM for robustness screening and to DAPP to inform trigger design. Where distributions are not credible, the flow bypasses this inner loop and proceeds directly to exploratory robustness analysis, preserving a decision-centred posture under deep uncertainty.

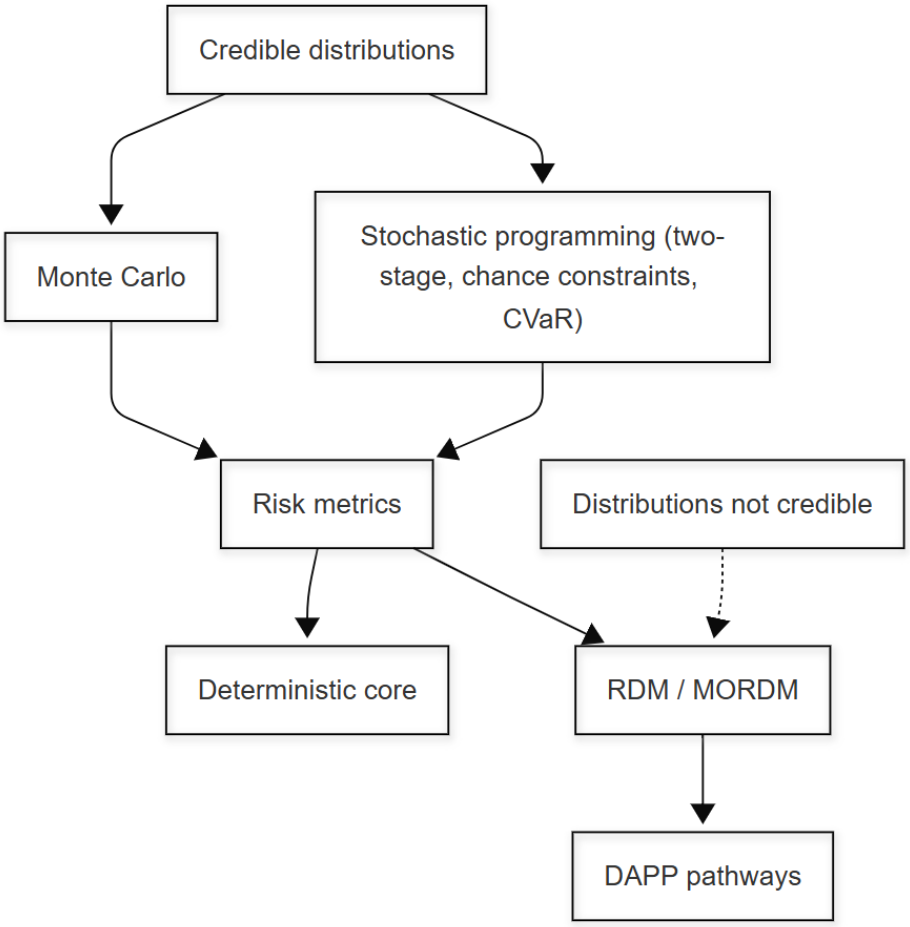


Figure 5.2 Probabilistic and stochastic inner loop (with L-shaped iteration)

However, their credibility depends critically on the existence, stability, and representativeness of the assumed probability distributions. For many epistemic or deep uncertainties, such as long-run carbon-price trajectories, the likelihood of disruptive technological shifts, long-term learning rates, or resource availability under climate and policy change, probabilities are often unknown, unknowable, or contested [9], [179]. In such cases, assumptions of stationarity, thin-tailed behaviour, or weak dependence may be violated, creating an illusion of precision that obscures exposure to structural and contextual drivers [310].

Consequently, while probabilistic and stochastic methods remain powerful for quantifying variability and operational risk in well-characterised domains, they are insufficient on their own for strategy design under deep uncertainty. In industrial process heat decarbonisation, they serve best as inner-loop tools, to analyse operational variability and short-term hedging, within a broader modelling framework that also employs exploratory and robustness-oriented approaches when probability assignments are uncertain or contested [9], [10], [50].

5.5 Decision-Making under Deep Uncertainty (DMDU)

Deep uncertainty denotes situations in which analysts cannot credibly specify a single system model, assign reliable probability distributions to key drivers, or agree on the values by which outcomes should be judged [9], [298]. Typical cases include unknown or disputed probabilities, data that are not informative about future consequences, and expert disagreement over policy effects [298]. In such settings, common for long-lived, capital-intensive process heat assets exposed to ETS policy credibility, delivered-biomass logistics, and GXP reinforcement timing, the aim shifts from prediction to robustness and adaptiveness; favouring strategies that meet agreed performance thresholds and avoid unacceptable outcomes across many plausible futures [18].

The DMDU paradigm formalises this stance by structuring problems with the **XLRM** schema, with e**X**ogenous uncertainties (X), policy **L**evers (L), system **R**elationships (R), and performance **M**etrics (M), and deploys exploratory modelling to stress-test candidate strategies over ensembles, diagnose vulnerabilities, and pair initial actions with contingent adaptations guided by signposts and triggers. Models are thus used as decision laboratories within a deliberation-with-analysis process, with evidence reported through transparent yardsticks such as satisficing coverage, regret (including tail quantiles), and reliability [9], [10].

This section reviews three core pillars and their immediate complements:

- **Robust Decision Making (RDM)** provides the ensemble-based evaluation harness and defines the robustness metrics used later in this thesis.
- **Dynamic Adaptive Policy Pathways (DAPP)** turns static plans into pathways of contingent actions, operationalised by observable signposts and explicit trigger thresholds.
- **Many-Objective RDM (MORDM)** integrates multi-objective optimisation (e.g., cost–emissions–reliability) with robustness screening to structure trade-offs transparently.

A short review of complementary building blocks, including Info-Gap, Decision Scaling, and Engineering Options Analysis, highlights how studies can be built from common elements rather than adopting every formalism wholesale.

5.5.1 Robust Decision Making (RDM)

Robust Decision Making (RDM) is a decision-support approach for deep uncertainty that shifts emphasis from prediction to robustness: the aim is to identify strategies that satisfy agreed performance thresholds and avoid unacceptable outcomes across wide ranges of plausible futures,

rather than to optimise against a single forecast [18], [49], [259]. Models serve as decision laboratories to explore how candidate strategies behave when key drivers and selected structural assumptions are varied, and to learn where and why those strategies become non-credible [10], [137].

For process heat decarbonisation, this means treating policy settings (e.g., carbon-price bands), infrastructure timing (e.g., GXP reinforcement), technology evolution (e.g., heat-pump performance and costs), and resource logistics (e.g., delivered biomass) as externally uncertain. The key question is which strategy designs remain defensible across diverse combinations of those conditions. RDM has been widely applied in water resources, coastal and flood risk, and long-lived infrastructure planning where probabilities are disputed, and policies may shift [9], [262]. These lessons transfer naturally to industrial process heat, where market instruments, infrastructure lead times, and supply-chain frictions co-determine outcomes.

In practical terms, the analysis proceeds by specifying a small set of candidate strategies (e.g., electrification-first, biomass-first, or hybrid/sequenced portfolios with high-level timing rules). It then constructs a computational ensemble that spans salient exogenous drivers and a few structural variants (e.g., chronology versus representative-days, or alternative site–region coupling representations). For each strategy p and future $\omega \in \Omega$ (a realisation of exogenous drivers and structural switches), the model evaluates performance $J(p, \omega)$ (e.g., system cost, emissions, reliability indices) subject to technical and policy constraints. Diagnostics focus on robustness yardsticks, rather than expected optimality. A standard measure is regret,

$$r(p, \omega) = J\,(p, \omega) - \min_{p'} J\,(p', \omega),$$

summarised by $\max_{\omega} r$, $\mathbb{E}[r]$, and high percentile regret (e.g., P90). A complementary measure is the *satisficing rate*,

$$S_{\theta}(p) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \mathbf{1}\{J\,(p, \omega) \leq \theta\}$$

for a decision-relevant threshold θ reflecting a decision criterion (cost, emissions, or reliability).

Reliability is the share of futures with no technical violations and reports both regret and satisficing rate within a future and across futures. Within a future ω, reliability is:

$$\mathrm{rel}(p, \omega) = 1 - \frac{\sum_t u_t(p, \omega)}{\sum_t d_t(\omega)}$$

where $u_t(p, \omega) \geq 0$ denotes unmet heat at time t and $d_t(\omega)$ the demand. Across futures, aggregate as the *reliability share*

$$R_{\varepsilon}(p) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \mathbf{1}\{\mathrm{rel}(p, \omega) \geq 1 - \varepsilon\},$$

for a small tolerance ε; optionally report tail summaries (e.g., a P90 on NPV or a CVaR-style shortfall) at level α:

$$\mathrm{CVaR}_{\alpha}[J] = \min_{\eta} \left\{ \eta + \frac{1}{1-\alpha} \mathbb{E}_{\omega}[(J(x, \omega) - \eta)_{+}] \right\}$$

Strategies are compared on these yardsticks, and vulnerability sets are identified (the combinations of drivers and assumptions under which targets are missed) [10], [18], [313], [314]. The scenario-discovery, then distils these conditions into transparent, rule-like statements.

For clarity of communication (and consistent with best practice in RDM applications), reporting should emphasise regret CDFs/percentiles by strategy, satisficing heatmaps against one or two pivotal drivers (e.g., carbon price versus delivered biomass cost), concise reliability summaries, and a short narrative of the conditions of failure [9], [314]. This evidentiary package shows not only which strategies are more robust, but also why, and at what trade-off in other objectives. In process heat setting, candidate strategies defined at a policy/portfolio level (not yet at unit commitment granularity); ensembles that reflect credible ranges for policy, infrastructure, technology, and logistics; a small, transparent set of robustness metrics; and clear narratives of vulnerability that can be translated into signposts and trigger rules, thereby preparing the ground for adaptive pathways (DAPP) and many-objective screening (MORDM).

Applications to energy systems have grown alongside these methodological developments. Yue et al. [311] review approaches to uncertainty in energy system optimisation models and identify the limited treatment of deep uncertainty in mainstream practice, while Paredes-Vergara et al. [312] develop a framework that integrates deep uncertainty into national power-system planning, using RDM in the Chilean case. These studies address the electricity system at national scale; the present thesis applies the same logic at the interface between an industrial site and its regional grid connection.

RDM's strengths are balanced by practical limitations. Evaluating thousands of futures for multiple strategies can be computationally intensive, particularly with high spatial and temporal resolution and multi-sector coupling. Moreover, although RDM does not require probabilistic forecasts, it still depends on a disciplined uncertainty envelope; broad enough to include critical futures yet compact enough to be tractable. Finally, robust conclusions typically emerge from deliberation with analysis, iterating with stakeholders to refine performance thresholds, confirm what counts as failure, and ensure findings are actionable in institutional settings [9], [137]. To manage these issues in process heat modelling, a feasible, staged and data-informed workflow needs to combine pre-screening

(deterministic/probabilistic runs to focus the deep-uncertainty set), structured sampling (e.g., Latin hypercube or other space-filling designs; [218], [306]), hierarchical evaluation (site to region) to keep ensembles tractable, and iterative pruning or clustering of behaviourally similar futures [19]. This preserves exploratory breadth while controlling runtime, and it delivers the robustness metrics and vulnerability sets needed to design adaptive pathways for process heat decarbonisation in New Zealand.

5.5.2 Dynamic Adaptive Policy Pathways (DAPP)

Dynamic Adaptive Policy Pathways (DAPP) is a planning approach that turns static plans into time-sequenced pathways of contingent actions designed to remain effective under deep uncertainty. Rather than committing to a single forecast, it identifies near-term actions that work across many plausible futures, defines signposts (indicators to monitor), and specifies triggers (decision rules) for switching, accelerating, or pausing actions when performance would otherwise fall below agreed thresholds. The central concept is the adaptation tipping point (the state of the world at which a strategy no longer meets performance objectives) together with the set of alternative pathways that preserve feasibility after such tipping points occur [19], [50]. In DAPP, "policy" includes actions by public, private, and community actors (e.g., plant managers, network operators, government) rather than government alone [137].

For Ω denote to the set of futures (realisations of exogenous drivers and any structural switches), a pathway is a policy π that maps observed indicator histories $h_t = \{z_0, \dots, z_t\}$ to actions $a_t = \pi(h_t)$. Its performance is written $J(\pi, \omega)$ for future $\omega \in \Omega$, evaluated with the same metrics used for strategies in RDM (e.g., discounted cost, cumulative emissions, reliability). An adaptation tipping point for a plan (or subplan) occurs at the first time a monitored metric $M_t(\pi, \omega)$ breaches its limit θ:

$$t^{\mathrm{ATP}}(\pi, \omega) = \inf\{t: M_t(\,, \omega) \geq \theta\}\,.$$

Consistent with Exploratory Modelling and Analysis (EMA) and the wider DMDU literature, pathway design is informed by ensembles that span exogenous uncertainties, model structures, and policy options. Ensemble experiments reveal where actions cease to meet objectives and which follow-up actions can restore performance [9], [19]. In multi-objective settings, robustness is expressed as performance that remains acceptable across many plausible futures, often represented by regret-based or satisficing measures, while Pareto analysis clarifies the trade-offs among objectives [315].

The adaptive logic of DAPP can be formalised within the framework of Partially Observable Markov Decision Processes (POMDPs), which generalise classical sequential decision models to situations

where system states are not fully observable and must be inferred from noisy or partial indicators [316]. In a POMDP, the decision-maker maintains a belief state, a probability distribution over possible system states, reflecting uncertainty about the true condition of the world [317]. At each time step t, the decision-maker holds a belief $b_t(s)$ over states $s \in S$, selects an action $a_t \in A$, receives an observation $o_t \in O$, and updates the belief using Bayes' rule:

$$b_{t+1}(s') = \eta\, O(\, o_{t+1} \mid s', a_t\,) \sum_{s \in S} T(\, s' \mid s, a_t\,) b_t(s)$$

where $T(\, s' \mid s, a\,)$ and $O(\, o_{t+1} \mid s', a\,)$ are the transition and observation models, and η normalises probabilities. The goal is to find a policy $\pi^*(b)$ that maximises expected cumulative reward (or minimises cost):

$$\pi^* = \arg\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t R(s_t, a_t) \,|\pi, b_t\right],$$

with discount factor $\gamma \in (0,1)$ [318], [319].

This formulation parallels the signpost–trigger mechanism of DAPP, where actions are chosen based on beliefs updated through monitoring rather than on perfect knowledge of the state. In the process heat context, hidden states may include policy credibility (ETS trajectory), unobserved grid-upgrade progress, or evolving technology costs, while observations correspond to measurable indicators such as sustained ETS prices, reinforcement announcements, or supply-chain data.

Although solving full POMDPs is computationally demanding, the conceptual analogy highlights that adaptive pathways function as belief-updating policies. Practical implementations can approximate this logic through rolling-horizon recourse, trigger thresholds, or state-dependent heuristics within an RDM–DAPP workflow. The POMDP framework therefore provides the formal anchor for interpreting DAPP as a structured sequential-decision process under uncertainty.

Viewed broadly, DAPP can be embedded within a hierarchy of learning-based decision models. If system states are fully observable, the problem reduces to a Markov Decision Process (MDP), in which actions directly update known states through transition probabilities [320]. In this setting, Reinforcement Learning (RL) algorithms (such as Q-learning [321] or policy-gradient methods [322]) can approximate optimal adaptive policies through simulation or real-time feedback, offering a potential future extension for dynamic plant operation or regional policy evaluation under changing conditions.

When state uncertainty is significant but can be progressively reduced through data, Bayesian networks provide a complementary representation; they encode probabilistic dependencies among uncertain variables (e.g., technology performance, policy credibility, resource availability) and can update beliefs as new evidence arrives [323]. Bayesian updating thus acts as a computational analogue of signpost monitoring, supporting probabilistic trigger calibration and early-warning detection of tipping points [324].

Together, these formulations (POMDP as the anchor, MDP/RL as learning extensions, and Bayesian networks as inference mechanisms) form a conceptual continuum linking DAPP's adaptive logic to dynamic optimisation and learning theory. While such methods are not implemented in the framework developing in the current study, they illustrate how adaptive pathways could evolve toward self-updating, data-driven policy systems as monitoring and computational capacity improve.

To assess pathway robustness quantitatively, two complementary measures are used: path-regret and the satisficing fraction.

The path-regret for policy π in future ω is:

$$r(\pi, \omega) = J(\pi, \omega) - \min_{\pi'} J(\pi', \omega),$$

summarised by maxima, means or high percentiles. The satisficing fraction measures how often a pathway meets defined performance thresholds:

$$S_\theta(\pi) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \mathbf{1}\{J(p, \pi) \leq \theta\},$$

for decision-relevant threshold θ (cost, emissions, reliability) [325].

Together, these indicators quantify how well a pathway maintains acceptable performance across futures and help prioritise those that combine high satisficing rates with low regret.

In practice, a DAPP study should:

- Define objectives and minimum performance thresholds (e.g., cost ceilings, emissions budgets, reliability targets);
- Use exploratory analysis (often RDM outputs) to identify vulnerability sets and prospective tipping points;
- Catalogue feasible actions and option-preserving steps with indicative lead times, costs, and prerequisites options;

- Design pathway maps that connect near-term actions to alternative follow-ups under different observed conditions; and
- Specify a monitoring and governance plan that assigns indicator ownership, review cadence, and decision rights, with explicit trigger thresholds, persistence windows (to avoid reacting to noise), and mobilisation lead-time allowances (time required to mobilise the next action) [50], [179].

For a signpost component $z_t^{(k)}$, a trigger is parameterised by a level θ_k, persistence window W_k, and lead time L_k. With a smoothed indicator $\bar{z}_{t,W_k}^{(k)}$, the trigger fires at

$$\tau_k(\omega) = \inf\left\{t \geq L_k : \bar{z}_{t,W_k}^{(k)}(\omega) \geq \theta_k\right\}$$

upon which the pre-defined contingent action is executed.

These formal definitions connect directly to the POMDP interpretation: triggers act as heuristic, belief-based decision rules that approximate Bayesian updating under partial observability.

Standard DAPP artefacts comprise a visual pathway map (with “transfer stations” and tipping points) and a scorecard that summarises costs, impacts, co-benefits, and risks for each pathway [50], [326]. The approach operationalises adaptivity in a way that is auditable and communicable. Pathway maps make contingent choices and potential lock-ins explicit; signposts and triggers reduce ad-hoc reversals; and lead-time accounting foregrounds practical feasibility. DAPP is complementary to RDM that ensemble stress-tests reveal where a plan fails; DAPP converts those insights into if-this-then-that rules and sequenced actions [50], [315]. Its limitations are largely organisational rather than mathematical: it requires institutional capacity to monitor indicators and act on triggers; poorly calibrated triggers can cause “churn” (too tight) or slow response (too loose); and early choices can create path dependence if options are closed prematurely [9]. Because many impact models otherwise treat adaptation as instantaneous and unconstrained, DAPP applications should represent practical constraints (biophysical, social, financial), time-lags and uptake limits, and actor-specific decision rules; triggers should, where possible, be objective and performance-linked, not modeller-chosen artefacts of the state representation [315], [327].

For industrial process heat, credible pathway design should consider:

(a) **Signposts:** ETS price bands; delivered-biomass costs within a defined logistics radius (including sustainability metrics); GXP hosting capacity and reinforcement milestones; performance/cost learning of high-temperature heat pumps; and site demand or product-mix changes.

(b) **Triggers:** combinations of threshold, persistence, and lead-time, for example, if the ETS price remains above a threshold P_{ETS}^{thr} for Y months and the local seasonal $\mathrm{COP} \geq \mathrm{COP}^{thr}$ with spare GXP headroom $\geq H_{GXP}^{thr}$, accelerate electrification; if delivered biomass cost $\leq C_{bio}^{thr}$ NZ$/MWh within a logistics radius R_{log} and sustainability criteria are met, prioritise biomass conversion; upon commissioning of network upgrade U by year T_U (increase in hosting capacity $\Delta H \geq \Delta H^{thr}$), enable electric-boiler/HTHP options at the identified sites.

(c) **Design elements:** no-regrets efficiency measures, option-preserving investments (e.g., connection foundations sized for later upsizing), and branch points aligned with policy or infrastructure review cycles [50], [326].

Where timing matters, time-evolving ensembles should be used to locate tipping points and the timing of transfers between actions, and pathway scorecards should report both near-term feasibility and long-term robustness [137], [315]. Internationally, DAPP has underpinned large-scale, multi-stakeholder transitions—e.g., the Dutch Delta Programme for flood risk management [50] and has been adapted for infrastructure and climate governance in other jurisdictions [328]. In the New Zealand context, DAPP helps structure investment timing (e.g., retire coal now vs defer until an ETS band is sustained or a GXP upgrade is commissioned) and technology switching (biomass-first vs electrification-first) by translating ensemble-based vulnerabilities into clear signposts, tying triggers to observable indicators and procurement lead times, and mapping contingent pathways that keep options open in regions where either biomass logistics or grid capacity could bind. The result is an adaptive and time-consistent strategy that remains credible across policy swings and infrastructure-timing uncertainty, while preserving the ability to pivot as information accrues [9], [326].

5.5.3 Many-Objective Robust Decision Making (MORDM)

Many-Objective Robust Decision Making (MORDM) integrates multi-objective optimisation with the robustness logic of Decision-Making under Deep Uncertainty (DMDU). Rather than seeking a single optimum, it explores a Pareto set of strategy designs across competing objectives (e.g., total cost, lifecycle emissions, reliability/security-of-supply, stranded-asset risk) and then applies robustness screening over large ensembles of plausible futures. The result is a robust-Pareto subset: designs that expose transparent trade-offs and remain defensible under deep uncertainty [9], [51].

MORDM typically employs Evolutionary Multi-Objective Optimisation (EMO) algorithms, such as NSGA-II, MOEA/D, or Borg, to approximate Pareto frontiers when analytical gradients are unavailable and the decision space is complex, discontinuous, or non-convex [314]. EMO methods evolve a

population of candidate solutions over successive generations, combining selection, crossover, and mutation operators to reveal the trade-off structure among competing objectives while preserving diversity across the front [329].

For $p \in \mathcal{P}$ denote a strategy design (decision-lever setting), $\omega \in \Omega$ a future (realisation of exogenous drivers and structural switches), and $J(p, \omega) = (J_1, \ldots, J_m)$ the vector of m objectives (to be minimised). A typical MORDM study proceeds in two stages:

1. Search many-objective evolutionary algorithm (MOEA): Use a many-objective evolutionary algorithm (e.g., Borg, ε-NSGA-II or NSGA-III) to approximate a Pareto set under a nominal evaluation (either a single reference future ω_0 or a small reference subset $\Omega_{\text{ref}} \in \Omega$). Concretely, solve

$$\min_{p \in P} F(p) = \begin{cases} J_i\,(p, \omega_0), & \text{single reference future,} \\ \dfrac{1}{|\Omega_{\text{ref}}|} \displaystyle\sum_{\Omega_{\text{ref}} \in \omega} J_i\,(p, \omega), & \text{reference subset,} \end{cases}$$

   Subject to feasibility constraints (technical/policy). The MOEA returns an approximate non-dominated set $\widehat{\mathcal{P}^*_{\text{ref}}}$.

2. Robustness screening (ensemble): Re-evaluate each $p \in \widehat{\mathcal{P}^*_{\text{ref}}}$ across the full ensemble Ω and retain designs that are robust on agreed yardsticks:

   - *Component-wise regret* (per objective *j*):

$$r_j(p, \omega) = J_i(p, \omega) - \min_{p' \in \widehat{\mathcal{P}^*_{\text{ref}}}} J_i\,(p', \omega),$$

     reported via $max_{\omega \in \Omega}\, r_j$, $\text{P90}_{\omega}[r_j]$, or the Cumulative Distribution Function (CDF) by j.

   - *Satisficing fraction* (non-probabilistic, fraction of futures):

$$S_\theta(p) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \mathbf{1}\left\{J_i(p, \omega) \le \theta_j \quad \forall j \in \mathcal{J}_{\text{sat}}\right\},$$

     for decision-relevant thresholds $\theta = (\theta_j)$.

   - *Reliability across futures*: use the reliability share $R_\varepsilon(p)$ defined earlier (fraction of futures with within-future reliability $\ge 1 - \varepsilon$).

     The robust-Pareto subset is then

$$\mathcal{P}_{\text{rob}} = \left\{p \in \widehat{\mathcal{P}^*_{\text{ref}}} : \text{stat}_{\omega}[r_j(p, \omega)] \le \tau_j \quad \forall j \in \mathcal{J}_{\text{reg}}, S_\theta(p) \ge \sigma, \text{Reliab}(p) \ge \rho\right\}$$

where $\mathrm{stat}_{\omega}[\cdot]$ is the chosen regret summary (e.g., max or P90 over Ω), and τ, σ, ρ are agreed screening cut-offs.

Beyond ensemble-based robustness screening, Multi-Attribute Utility under Uncertainty offers a complementary, value-analytic framing. Robustness metrics, such as regret, satisficing rate, or reliability, can themselves be treated as attributes in a multi-attribute utility function, integrated with conventional performance objectives (cost, emissions, flexibility) [330]. This enables value-based MCDA under uncertainty, where stakeholder preferences, trade-offs, and tolerance to risk or regret are explicitly represented [331]. In this sense, MORDM extends beyond optimisation to a decision-analytic synthesis, combining quantitative robustness with qualitative stakeholder values within a unified framework

In industrial process heat decarbonisation, a MORDM study specifies objectives, constraints, and decision levers at the strategy level (e.g., portfolio shares, capacity bounds, conversion-timing windows, or option-preserving investments). The algorithm generates a candidate Pareto set, which is then re-evaluated across ensembles spanning exogenous drivers and selected structural variants [314], [329].

Designs are retained if they:

(i) meet satisficing thresholds (e.g., cost $\leq \theta_1$, emissions $\leq \theta_2$, unserved-heat $\leq \theta_3$) in a large fraction of futures;

(ii) exhibit low regret (e.g., max or P90 below bounds); and

(iii) maintain reliability (no technical violations)

The MOEA reveals structure in the trade space (front curvature, “knees”), while the robustness filter removes designs that depend on narrow, fragile assumptions. The surviving designs form a shortlist for pathway construction (DAPP) or, if required, stakeholder weighting through MCDA.

Reporting should emphasise:

- standard Pareto visualisations (scatter/parallel-coordinate plots) showing the diversity of trade-offs;
- robustness overlays, including satisficing coverage, regret percentiles, and reliability bars; and
- concise vulnerability narratives that identify the combinations of drivers undermining performance, linking forward to scenario discovery.

MORDM thus exposes decision-relevant trade-offs while avoiding spurious precision about probabilities and naturally supports deliberation since no single metric dominates. Challenges include computational burden (many evaluations), dimensionality (too many objectives obscure structure), and sample sensitivity (robustness scores depend on the ensemble). Practical solutions include staged searches (coarse-to-refined), space-filling ensembles for robustness checks, and surrogate or emulator models to accelerate evaluation [325]. As noted in energy modelling reviews, hybrid approaches with surrogates can mitigate these demands in complex systems like heat decarbonisation [202], [332].

In the context of decarbonising process heat, typical objectives include discounted system cost, cumulative CO_2-e, resource security (exposure to biomass scarcity or grid constraints), and transition risk (e.g., stranded-asset write-downs). Constraints reflect policy and physical limits; ETS budgets/bands, GXP hosting capacity, regional biomass sustainability/logistics, and site-level operational feasibility. In this setting, MORDM (i) identifies knee-point designs that balance cost and emissions while maintaining reliability under GXP/biomass limits, (ii) highlights designs that remain robust across carbon-price bands and reinforcement lead times, and (iii) clarifies conditions of failure that motivate DAPP triggers (e.g., price bands, biomass-cost corridors, commissioning milestones) [51], [314]. These elements collectively underpin the robust and adaptive modelling framework implemented in this thesis.

5.5.4 Complementary DMDU building blocks

Recent DMDU guidance emphasises composing analyses from common building blocks rather than adopting every formalism wholesale. A generic loop provides a scaffold: frame → explore → choose initial and contingent actions → monitor and iterate [9], [50]. Three complementary building blocks are Info-Gap, Decision Scaling, and Engineering Options Analysis, which are particularly useful alongside RDM/DAPP/MORDM for industrial process heat.

- **Info-Gap (IG) decision theory** treats deep uncertainty as an expanding "horizon of unknowns" around a nominal driver state and asks how much deviation a strategy can tolerate while still meeting thresholds θ. For a nominal vector of exogenous drivers $x_0 \in \Xi$ (e.g., ETS, delivered biomass cost, grid CO_2 intensity), and $\mathcal{U}(\alpha;\ x_0) \subset \Xi$ an uncertainty set of radius α around a nominal condition x_0. For a scalar performance threshold θ (or component-wise vector), the robustness function is:

$$\hat{\alpha}_\theta(p) = \sup\{\alpha : J(p,u) \leq \theta \quad \forall\ u \in \mathcal{U}(\alpha;\ x_0)\}.$$

An “opportuneness” analogue can be defined as

$$\hat{\alpha}_{\phi}^{+}(p) = \inf\{\alpha : \exists\, x \in \mathcal{U}(\alpha;\, x_0), \quad with\, J(p,u) \le \phi\}.$$

IG is attractive when probability distributions are not defensible, but performance thresholds are clear (e.g., reliability $\geq R^*$, emissions $\leq E^*$) [333], [334], noting sensitivities to the nominal point and uncertainty norm [335]. In this thesis, IG is used as a reporting layer on the RDM ensemble by sweeping one driver (e.g., ETS, delivered-biomass cost, grid CO_2) around the discovered boundary to trace $\hat{\alpha}_{\theta}(p)$ rather than a separate formalism.

- **Decision Scaling (DS)** starts from performance requirements and maps where a strategy fails in the exogenous-driver space. For $x \in \Xi$ denote a vector of drivers (prices, demand growth, weather, infrastructure timing). Define the safe operating set

$$S_{\theta}(p) = \{x \in \Xi : J(p,x) \le \theta\}, \qquad \mathcal{V}_{\theta}(p) = \Xi \setminus S_{\theta}(p).$$

On a finite ensemble Ω, the empirical set is

$$\hat{S}_{\theta}(p) = \{\omega \in \Omega : J(p,\omega) \le \theta\},$$

with the fraction-of-futures satisficing measure $S_{\theta}(p) = \frac{|\hat{S}_{\theta}(p)|}{|\Omega|}$. DS complements RDM by focusing attention on $\mathcal{V}_{\theta}(p)$ before connecting to external scenarios; in practice we deliver DS via the same experiments as RDM plus a vulnerability map (often aided by PRIM/trees) [9], [336].

- **Engineering Options Analysis (EOA)** values flexibility embedded in design (staging, deferral, expansion, switching) because timing choices have real economic value under uncertainty [337]. Without committing to a probabilistic real-options model, a non-probabilistic flexibility value can be expressed as the gain in robustness when augmenting a base strategy p with a contingent action set $\mathcal{A}$ (e.g., staged capacity, dual-fuel retention):

$$\Delta S_{\theta}(p;\mathcal{A}) = S_{\theta}(p \oplus \mathcal{A}) - S_{\theta}(p),$$

$$\Delta\mathrm{Regret}(p;\mathcal{A}) = stat_{\omega\in\Omega}\,[r(p,\omega)] - stat_{\omega\in\Omega}\,[r(p \oplus \mathcal{A},\omega)],$$

where $S_{\theta}(\cdot)$ is the *fraction-of-futures* satisficing measure, $r(p,\omega)$ is regret and $stat_{\omega\in\Omega}[\cdot]$ is a chosen regret summary (e.g., max or P90), and $p \oplus \mathcal{A}$ denotes “p with contingent actions enabled”. Positive ΔS_{θ} and $\Delta\mathrm{Regret}$ indicate beneficial flexibility. In this thesis, EOA is reflected

operationally through DAPP triggers and rolling-horizon solves; full real-options valuation (e.g., least-squares Monte Carlo) is reserved for future work where data and institutional needs warrant it [9], [338].

To situate these tools within the overall analysis, a compact mapping is provided in Table 5.2. Each building block is positioned by its guiding question, the circumstances under which it is most informative, the lightweight artefact adopted here, and its point of integration with the RDM–DAPP–MORDM workflow.

Table 5-2 Complementary DMDU Building Blocks for Process Heat Decarbonisation

Building Block	Central question	When it helps	Artefact in this thesis	Where it plugs in
Info-Gap	How much surprise can this strategy tolerate and still meet thresholds?	Thresholds are clear; probabilities are contested	Robustness curve vs ETS/biomass/electricity-CO_2 intensity	Reporting layer on the RDM ensemble
Decision Scaling	Where are the failure regions before applying probabilities?	Early framing; stakeholder dialogue on "what breaks"	Vulnerability map in driver space; safe operating regions	Scenario Discovery
EOA	What is the option value of staging/deferral/switching?	Lumpy CAPEX; timing choices matter	Pathway with staged actions; note option value qualitatively	Implemented via DAPP triggers & rolling-horizon

Robust frameworks are computation- and process-heavy. In prototypes, widening the uncertainty envelope and reporting clear robustness and vulnerability metrics is often more valuable than adding another full formalism [9]. Accordingly, IG, DS, and EOA are reviewed for completeness, DS-style vulnerability mapping and EOA-type flexibility are implemented via DAPP and rolling horizons, and full IG and real-options treatments are reserved for future extensions. Taken together, these elements focus attention on the systematic surfacing of the conditions under which otherwise promising strategies cease to satisfy standards and the clear communication of those conditions so they can be guarded with triggers or redesigned.

Figure 5.3 synthesises the DMDU workflow adopted in this thesis. XLRM framing defines uncertainties, levers, relationships, and metrics; candidate strategies are stress-tested over an ensemble; robustness metrics guide scenario discovery of vulnerability rules and, where appropriate, many-objective screening; these insights are translated into adaptive pathways with signposts and triggers, closing the loop via deliberation and iteration.

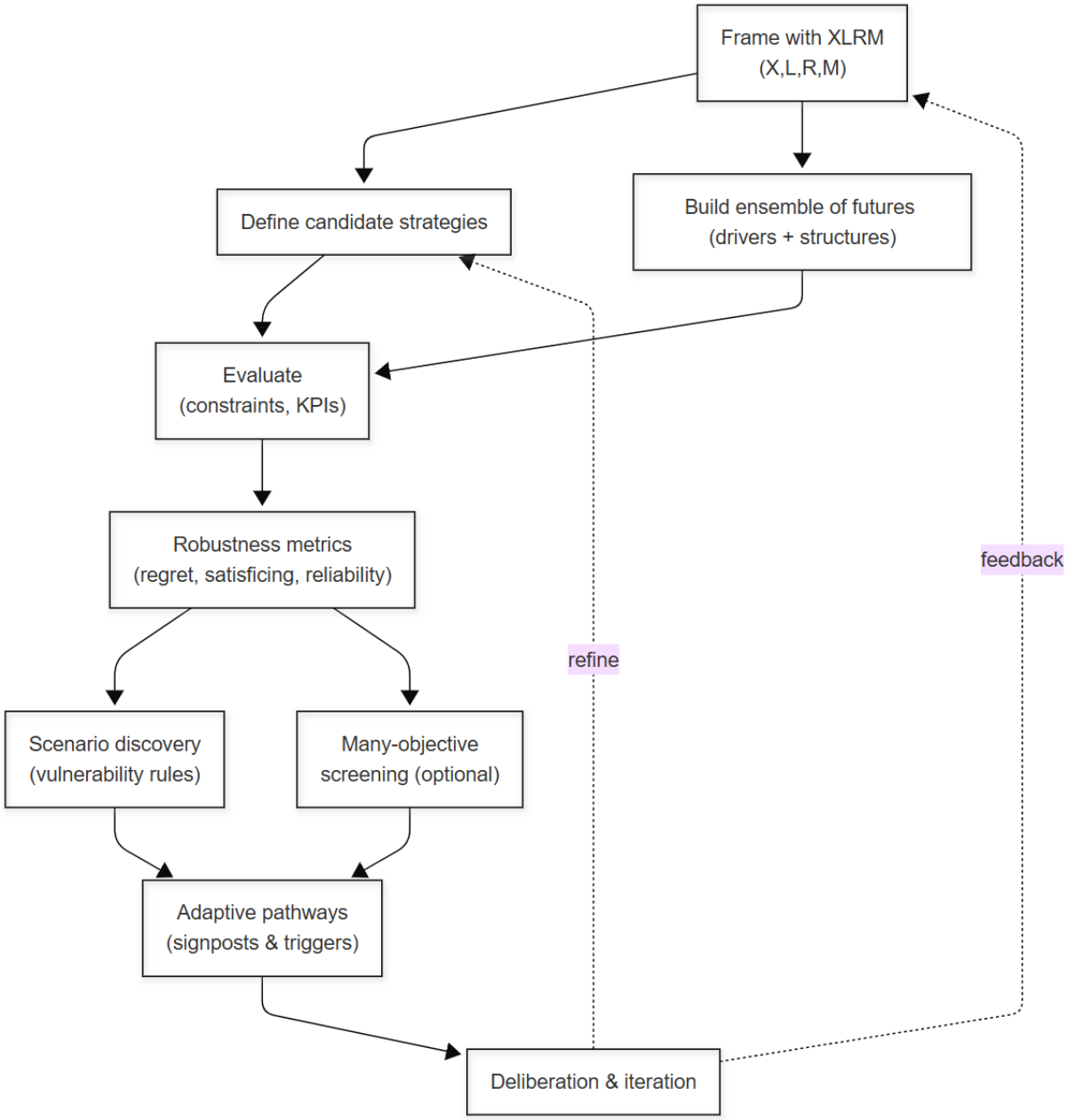


Figure 5.3 Conceptual DMDU workflow for process heat decarbonisation

5.6 Scenario discovery and backcasting

Under deep uncertainty, large ensembles from RDM/MORDM are most useful when their implications can be conveyed as simple, defensible conditions for success and failure. Scenario discovery addresses this need by extracting simple, human-readable *vulnerability sets*; rules that describe the combinations of exogenous drivers under which a strategy fails agreed thresholds (e.g., total cost, emissions, reliability) [9], [261]. The aim is explanatory parsimony rather than prediction: the central

question becomes, under what ranges of ETS price, delivered-biomass cost/radius, GXP headroom and timing, demand growth, or technology performance does strategy succeed or fail?

For Ω denote the ensemble of futures and $m(p,\omega) \in \mathbb{R}^K$ collect K performance metrics (e.g., cost, emissions, reliability indices) for strategy p in future ω. With thresholds $\theta \in \mathbb{R}^K$ and a sign vector $s \in \{+1,-1\}^K$ (so that each criterion is converted to "smaller is better" via $s_k m_k$), define a pass/fail indicator

$$y(p,\omega) = \mathbf{1}\{\, s_k m_k(p,\omega) \le s_k \theta_k \;\; \forall k = 1,\dots,K \,\}, \qquad \omega \in \Omega.$$

and the failure set $F(p) = \{\, \omega \in \Omega : y(p,\omega) = 0 \,\}$, and the driver map $x : \Omega \to \mathcal{D} \subset \mathbb{R}^d$ that extracts the d- dimensional driver vector (e.g., ETS, biomass cost, headroom).

An operational workflow (i) defines a binary outcome for each strategy–future run (meets all thresholds vs fails), (ii) selects a concise and interpretable set of drivers (policy bands, logistics, infrastructure milestones, technology parameters, demand/weather), and (iii) induces low-complexity rules that balance coverage/support (share of failing cases captured) and precision/density (share of cases inside the rule that truly fail). The Patient Rule Induction Method (PRIM) identifies high-failure "boxes" in driver space via patient peel–paste heuristics, returning interpretable hyper-rectangles

$$B = \prod_j [\, l_j, u_j \,] \subset \mathcal{D}$$

via patient peel–paste heuristics (e.g., "if ETS $\in [P_1, P_2]$ and biomass $\le C^*$ and headroom $< H^*$ then fail) [249], [339]. For any candidate box B, let $\Omega(B) = \{\omega \in \Omega : x(\omega) \in B\}$.

Then coverage /recall

$$\kappa(B) = \frac{|\, F(p) \cap B \,|}{|\, F(p) \,|} \;,$$

precision/density

$$\delta(B) = \frac{|\, F(p) \cap B \,|}{|\, \Omega \cap B \,|} \;,$$

with support

$$\sigma(B) = \frac{|\, \Omega \cap B \,|}{|\, \Omega \,|}$$

Shallow decision trees (CART) offer a complementary, visual rule set and driver ranking [340]. Good practice reports rule quality $\{k, \delta, \sigma\}$, stability under cross-validation, and clear visuals (rule boxes on

2-D driver maps, peeling trajectories, partial-dependence/accumulated-effects plots, and confusion matrices) to communicate directionality and error trade-offs [340], [341].

Backcasting proceeds from endpoints to actions, given performance targets and deadlines, what minimal sequence of decisions preserves feasibility once lead times and uncertainty are acknowledged? For time indexed by $t \in \mathcal{T}$, time-profiled thresholds θ_t, and per-period metrics $m_t(p, \omega)$, the adaptation tipping time

$$\tau(p, \omega) = \inf \left\{ t \in \mathcal{T} : \exists\, k\ s.t. \quad s_k m_{t,k}(p, \omega) \geq s_k \theta_{t,k} \right\}$$

for example, the earliest time at which the current plan would cease to meet objectives. Pathway maps are then constructed that connect near-term "no-regrets" steps to contingent follow-ups conditioned on monitored indicators (signposts) and explicit trigger logic. For an observable indicator stream $z_t(\omega)$ (e.g., ETS band, delivered-biomass cost, confirmed GXP upgrade date). A trigger is a decision rule $\mathcal{T}(z_{t-W:t}) \in \{0, 1\}$ that fires when an indicator condition holds persistently over a window W and schedules a contingent action after a lead time L. In multi-objective settings, backcasting is typically informed by robust-satisficing and regret-based yardsticks from the preceding ensemble analysis and is paired with pathway scorecards that summarise costs, co-benefits, and risks per branch [9], [50]. The notion of adaptation tipping points organises when branches should be taken [253], [254].

Evidence from water resources, coastal protection, and broader climate-risk planning shows that participatory, computer-assisted discovery improves transparency and uptake, particularly when PRIM or shallow trees are used to co-produce vulnerability statements that later anchor signposts and trigger thresholds [249], [341]. Discovery is commonly embedded within DMDU workflows to diagnose failure regions, prioritise monitoring indicators, and target learning; backcasting then sequences options consistent with institutional lead times and budget cycles [9], [261].

In process heat decarbonisation, salient drivers include ETS price bands/paths, delivered-biomass costs within a logistics radius (with sustainability status), GXP hosting capacity and reinforcement milestones, technology parameters (e.g., high-temperature heat-pump COP and CAPEX), demand growth, and selected structural switches (e.g., chronology vs representative-day modelling). Scenario discovery would be expected to yield sparse, auditable rules such as "Electrify-first fails when ETS$\leq P_1$ and local GXP headroom$\leq H^*$ prior to upgrade T^*, unless COP$\geq k$", or "Biomass-first fails when delivered biomass$\geq C^*$ within Z km or regional sustainability flags are unmet", each reported with $\{k, \delta, \sigma\}$ [9], [249]. Backcasting would then identify the minimal action set that preserves feasibility (e.g., complete efficiency package E by t_1; secure connection capacity ≥ H* by t_2; place a conditional

electrification order if ETS $\geq P^*$ for $\geq Y$ Y months; procure a biomass tranche if delivered cost $\leq C^*$ for two seasons) and align these with the signposts and trigger rules developed under DAPP [9], [50].

To maintain auditability and facilitate stakeholder deliberation, a compact artefact set is recommended: (i) a one-page vulnerability map per strategy (rule statements with coverage/precision and validation stats); (ii) phase diagrams for key driver pairs with rule boxes overlaid; and (iii) a backcasting corridor that links mandated endpoints to near-term actions, signposts, and triggers. Provenance should be documented through an ensemble manifest (driver ranges, sampling design), solver/code versions, and acceptance tests (feasibility, reliability, emissions thresholds) [9], [249], [341].

Figure 5.4 summarises the pipeline from ensemble stress-testing to adaptive pathways. Each strategy–future run in the ensemble Ω is labelled pass/fail against thresholds θ. A concise driver set $x(\omega) \in \mathcal{D}$ is selected, and interpretable rules are induced via PRIM (hyper-rectangles or "boxes") and shallow CART (visual splits and driver ranking). Rule quality is reported with coverage k, precision δ, and support σ, with cross-validation for stability. These auditable vulnerability statements then anchor backcasting: endpoints and deadlines are translated into minimal action sequences with signposts and trigger logic (DAPP). Monitoring closes the loop, feeding new evidence back to the ensemble design and rule checks.

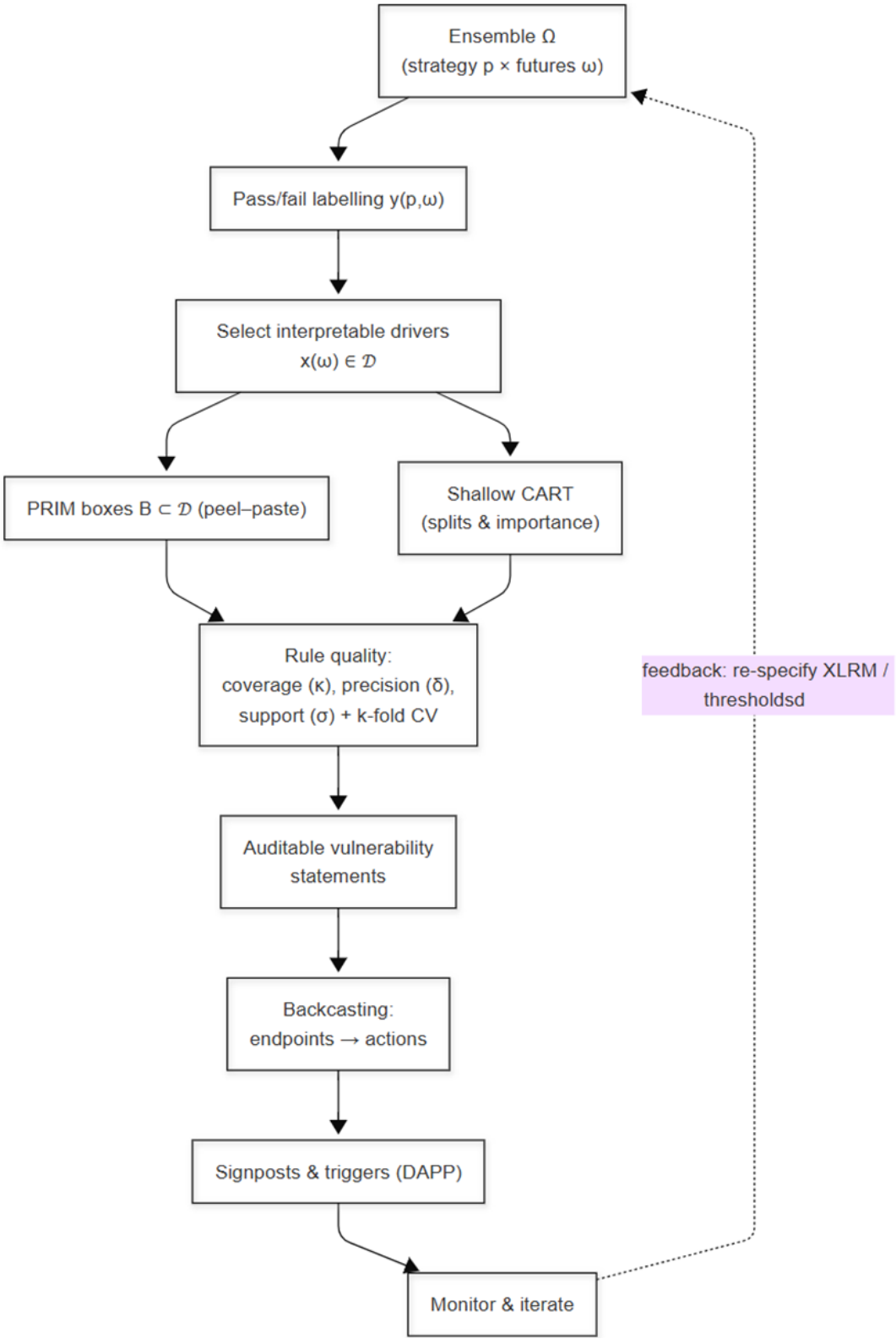


Figure 5.4 Conceptual pipeline for scenario discovery and backcasting

Discovered rules are contingent on the feature set and ensemble design; over-peeled boxes or deep trees can erode interpretability, while too-narrow ensembles may miss critical failure modes. Emphasis should therefore be placed on sparse rules, documented stability (hold-out or k-fold), and direct

alignment between predicates and observable indicators to support monitoring and governance [9], [341]. Computational demands are moderate for interpretable methods (PRIM, shallow CART) but grow with ensemble size; hybrid approaches (e.g., surrogate models for boundary re-checks) are increasingly used in high-resolution energy studies [314], [329].

In keeping with a small-strategy/large-worlds DMDU design for New Zealand process heat, scenario discovery is performed per strategy using a parsimonious PRIM-lite box search and a shallow decision tree to produce interpretable vulnerability statements; comparative dominance maps then show where one strategy outperforms another on robustness. Boundary regions are oversampled, boundary cases are validated at full temporal resolution (and with electricity-system coupling as needed), and DAPP triggers are derived directly from the discovered thresholds. This preserves the core DMDU principles (ensemble stress-testing, robustness metrics, and adaptive pathways) while remaining computationally manageable and transparent for stakeholder review.

5.7 Multi-Criteria Decision Analysis (MCDA)

Robust and adaptive analyses clarify which strategy designs remain credible under deep uncertainty; however, real decisions must still reconcile plural, non-commensurate objectives, such as discounted system cost, lifecycle emissions, reliability and security of supply, local air-quality impacts, and regional employment, together with heterogeneous stakeholder values.

Multi-Criteria Decision Analysis (MCDA) provides a formal, transparent, and auditable layer to structure these trade-offs when criteria cannot (or should not) be fully monetised [342], [343]. In energy-system and sustainability studies, MCDA is used to rank, sort, or select technologies and policy portfolios while keeping environmental, economic, and social attributes visible [344]. The principal methodological families include:

- value/utility models (MAVT/MAUT, SMART/SMARTER, AHP) [345];
- outranking relations (ELECTRE, PROMETHEE) [346]; and
- rule-based and dominance-based approaches [347].

No single method is universally superior; “fit-for-purpose” selection should reflect the structure of trade-offs, data quality, and the desired degree of stakeholder participation [348].

A standard formulation begins with a robust shortlist P_{short} of candidate strategies and a criterion set J includes conventional indicators (discounted cost, lifecycle CO_2-e, peak MW, local air-quality proxies,

indicative employment) and robustness-derived metrics introduced earlier, such as regret percentiles $r_{\mathrm{P90}}(p)$, satisficing coverage $S_\theta(p)$, reliability $R_\varepsilon(p)$.

For each $p \in P_{\mathrm{short}}$ and $j \in J$, the analysis uses a performance score $y_j(p)$, which is mapped to a common value scale $v_i(y_j(p)) \in [0,1]$ with an explicit preference direction (decreasing for costs/emissions/regret; increasing for reliabilities or satisficing rates).

With non-negative criterion weights w_j summing to 1, a compensatory additive value is:

$$V(p) = \sum_{j \in J} w_j \; v_j\left(y_j(p)\right), \qquad p \in P_{\mathrm{short}}$$

used to rank strategies when independence and preferential-independence trade-off assumptions are acceptable [342].

In the process heat context, this matrix typically captures New Zealand-specific considerations such as ETS-driven cost exposure, delivered-biomass availability within a logistics radius (and associated sustainability status), and GXP hosting-capacity constraints influencing electrification reliability [348].

When limited or no trade-off is acceptable, such as minimum reliability or air-quality standards, non-compensatory checks are advised [342]. PROMETHEE implements an outranking relation via criterion-wise preference functions $P_j(\Delta) \in [0,1]$ (with optional calibrated indifference/preference thresholds) and reports a net-flow score:

$$\Phi(p) = \Phi^+(p) - \Phi^-(p)$$

$$\Phi^+(p) = \frac{1}{|P_{\mathrm{short}}| - 1} \sum_{\substack{q \in P_{\mathrm{short}} \\ q \neq p}} \sum_{j \in J} w_j \, P_j\left(y_j(p) - y_j(q)\right)$$

$$\Phi^-(p) = \frac{1}{|P_{\mathrm{short}}| - 1} \sum_{\substack{q \in P_{\mathrm{short}} \\ q \neq p}} \sum_{j \in J} w_j \, P_j\left(y_j(p) - y_j(q)\right)$$

Larger $\Phi(p)$ indicates broader, criterion-weighted preference for p over competitors. ELECTRE-style vetoes can enforce non-substitutability on must-meet criteria (e.g., reliability or local air-quality thresholds) [343], [346], [349].

Where assessments are ordinal or linguistic rather than cardinal, the Dominance-Based Rough Set Approach (DRSA) offers a transparent, rule-induction alternative to numerical aggregation [347]. DRSA derives decision rules consistent with stakeholder-supplied preference orderings, thereby identifying dominance and inconsistency patterns within the data. This is especially useful when sustainability

attributes or qualitative stakeholder preferences cannot be meaningfully quantified but should remain decision-visible.

For value modelling and weights, practical elicitation tools include swing-weighting/SMART(ER) and AHP (with consistency checks); MACBETH can be used to construct judgement-based value functions [350], [351], [352].

Weights and value-function parameters are themselves uncertain; therefore, rankings should include stability diagnostics rather than a single deterministic outcome. For an admissible weight set

$$\mathcal{W} \subseteq \{ w \in \mathbb{R}_+^{|J|} : \sum_j w_j = 1 \} ,$$

a compact robustness indicator is the acceptability index,

$$\pi(p) = \frac{1}{|\mathcal{W}|} |\{ w \in \mathcal{W} : V_w(p) \geq V_w(q),\ \forall\, q \in P_{\text{short}} \}|$$

, reported alongside weight-sweep plots; high $\pi(p)$ indicates rankings stable to plausible preference variation [343], [353]. For imprecise or interval-valued data, outranking or DRSA models enhance interpretability and legitimacy [343], [344].

Many industrial decisions are portfolio-based, involving multiple sites subject to shared budgets or policy constraints. MCDA scores can thus feed a budget-constrained selection problem that preserves system interactions:

$$\max_x \sum_{s \in S} x_s U_s \quad s.t. \quad \sum_{s \in S} x_s C_s \leq B \qquad Ax \leq b\ ,\ \ x_s \in \{0,1\}\ ,\ \ \forall\ s \in S$$

where U_s is the site-level MCDA value, C_s the investment cost, B the total budget, and $Ax \leq b$ encodes regional limits such as GXP headroom or regional biomass-sustainability caps. U_s may include robustness-derived columns (e.g., S_θ, r_{P90}, R_ε) to keep uncertainty signals visible in portfolio choice. Rankings can then be revisited as monitoring indicators cross signposts within adaptive-pathway cycles, allowing transparent, time-consistent updates.

Reporting should include:

- the performance matrix with robustness columns explicit,
- the baseline additive ranking;
- a non-compensatory cross-check (e.g., PROMETHEE/ELECTRE/DRSA); and

- ranking-robustness diagnostics, such as weight-sweep visuals and an acceptability index $\pi(p)$ for top-rank stability.

Potential issues, including rank reversal, threshold calibration, and weight sensitivity, are mitigated through anchored value scales, reference-level normalisation, global sensitivity analysis, and transparent parameter documentation [342], [354]. This posture (anchored scales, explicit uncertainty handling, and clear limits to compensation) aligns with the thesis's post-positivist stance that models are structured aids to deliberation rather than instruments of prescriptive certainty.

5.7.1 Value-of-Information (VoI) analysis

Value-of-Information (VoI) analysis provides a complementary bridge between model evaluation and real-world data priorities. It quantifies the expected benefit of reducing uncertainty in a given parameter, such as delivered biomass cost, network reinforcement timing, or process-efficiency factors, by measuring the improvement in decision quality (e.g., expected value or reduction in regret) if perfect or improved information were available about uncertain parameters [345].

Formally, for decision $a \in A$, uncertain future $\omega \in \Omega$, and utility (or net-benefit) function $U(a, \omega)$:

$$\mathrm{EVPI} = \mathbb{E}_{\omega}\left[\max_{a\in A} U\,(a, \omega)\right] - \max_{a\in A} \mathbb{E}_{\omega}\,[U(a, \omega)],$$

$$\mathrm{EVPPI}_{\theta_k} = \mathbb{E}_{\theta_k}\left[\max_{a\in A} \mathbb{E}_{\omega|\theta_k} U\,(a, \omega)\right] - \max_{a\in A} \mathbb{E}_{\omega}\,[U(a, \omega)],$$

$$\mathrm{EVSI} = \mathbb{E}_{D}\left[\max_{a\in A} \mathbb{E}_{\omega|D}\, U\,(a, \omega)\right] - \max_{a\in A} \mathbb{E}_{\omega}\,[U(a, \omega)],$$

Here EVPI (Expected Value of Perfect Information) gives the upper bound on the worth of perfect information, EVPPI (Expected Value of Partial Perfect Information) isolates the contribution of a subset of parameters θ_k, and EVSI (Expected Value of Sample Information) estimates the value of obtaining new imperfect data D [355].

In the process heat context, VoI highlights where targeted data collection or model refinement is economically justified, for instance, when uncertainty about delivered biomass cost, grid-reinforcement timing, or heat-pump performance drives major differences in portfolio choice. By comparing EVSI with information-acquisition costs, yielding the Expected Net Benefit of Sampling, ENBS=EVSI − Cost(data), VoI identifies whether further data gathering (e.g. sub-GXP grid modelling or high-resolution biomass supply mapping) provides sufficient expected decision improvement to warrant investment.

VoI therefore reframes uncertainty reduction as a decision-directed resource-allocation problem; information is itself a constrained resource. This operationalises the decision-first, roll-up-second principle central to the framework (formalised in Chapter 6 as the decision-first boundary principle), directing analytical and empirical effort toward uncertainties that demonstrably affect decision robustness, not toward abstract model completeness [356]. In practice, EVPI and EVPPI inform which parameters matter most; EVSI and ENBS guide where new data collection is worthwhile; and integration with robustness metrics (regret, satisficing coverage, reliability) ensures that model refinement remains proportionate to its decision leverage. Thus, VoI serves as a quantitative bridge between model evaluation and information governance, enabling iterative, evidence-based model improvement under resource constraints [345], [355].

5.7.2 Participatory MCDA and Co-production frameworks

Beyond technical ranking and VoI diagnostics, participatory and co-production frameworks embed stakeholders directly into the weighting, scenario, and pathway-development processes. Criteria, weights, and pathways become co-produced artefacts rather than modeller-imposed assumptions, enhancing procedural legitimacy and the real-world uptake of results [357].

Recent literature conceptualises co-production as an iterative cycle of co-commissioning, co-design, co-delivery, and co-assessment, each phase embedding learning and adjustment [358]. Integrating MCDA and VoI within this cycle aligns the analytical and participatory layers:

- criteria and value functions are co-commissioned through stakeholder consultation;
- weighting, scenario framing, and data-priority decisions are co-designed using VoI diagnostics;
- adaptive strategies are co-delivered via portfolio implementation; and
- performance monitoring and pathway revision form the co-assessment phase feeding back into subsequent cycles.

Figure 5.5 summarises the MCDA decision layer and its linkages to the wider framework. Robustness outputs from the ensemble analysis (regret, satisficing coverage, reliability) are combined with conventional criteria into a performance matrix, mapped to value scales, and synthesised via either an additive value model (MAVT/MAUT) or non-compensatory methods (PROMETHEE/ELECTRE/DRSA). The resulting rankings feed a budget-constrained portfolio selection that preserves system couplings ($Ax \leq b$), after which monitoring supports pathway revisions. Value-of-Information (EVPI/EVPPI/EVSI

and ENBS) guides data-collection priorities, while participatory co-production informs criteria, weights, and thresholds.

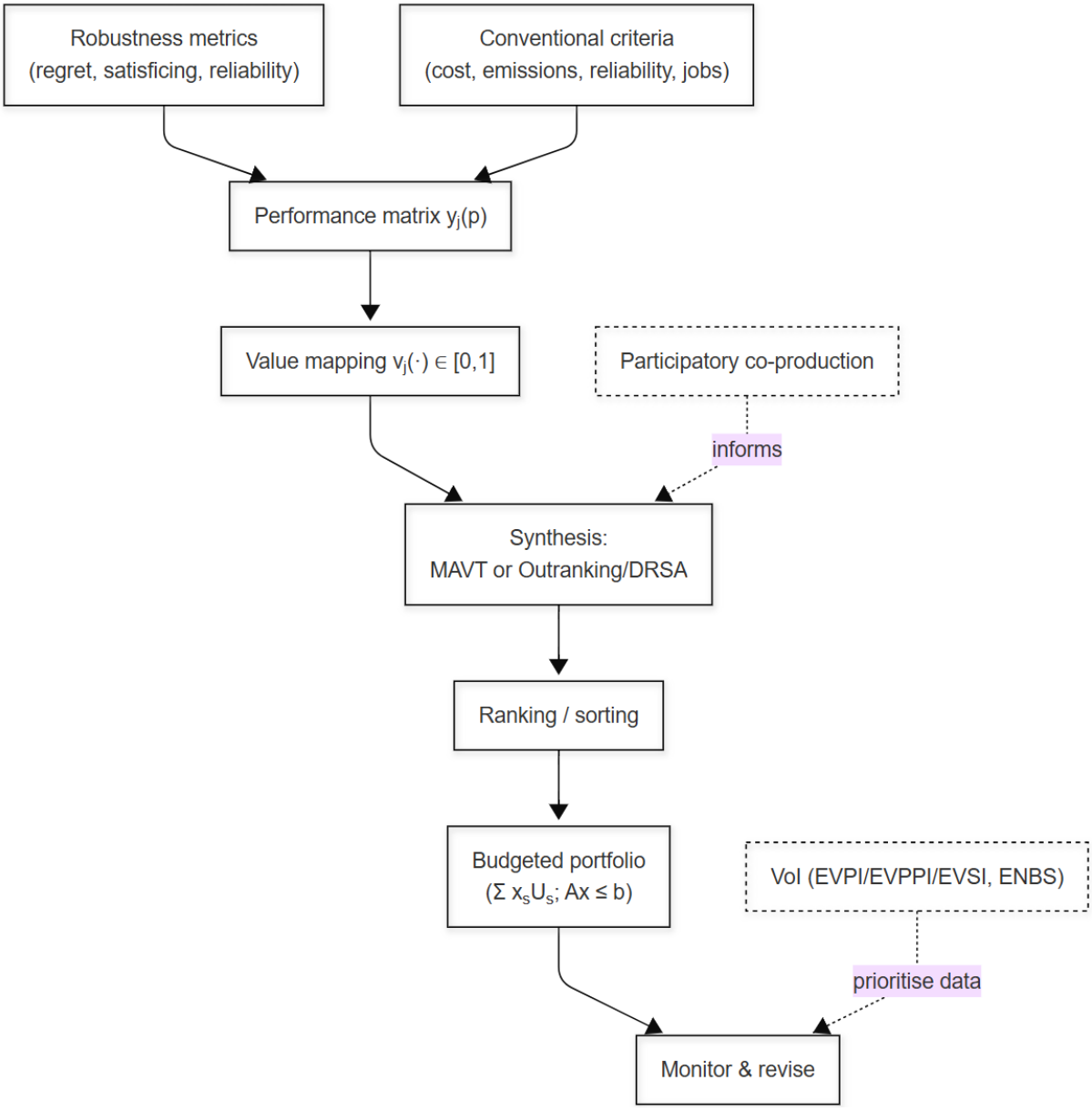


Figure 5.5 MCDA decision layer within the process heat framework

In New Zealand's polycentric energy-governance environment (spanning EECA, MBIE, regional councils, and industrial actors) such participatory embedding is essential for credible, legitimate transition planning. Structured workshops, transparent handling of value conflicts, and joint exploration of trade-offs ensure that MCDA and VoI outputs remain interpretable, contestable, and actionable within adaptive-planning horizons.

5.8 Socio-technical and behavioural extensions

Robust, ensemble-based evaluation reveals where strategies succeed or fail; dynamic and behavioural models help explain how transitions unfold over time and why actors may not follow frictionless optimisation. While the deterministic–exploratory core of the framework addresses technical feasibility and robustness, socio-technical extensions such as System Dynamics (SD) and Agent-Based Modelling (ABM) operate as outer-layer modules. These components do not alter the optimisation logic itself but provide dynamic or behavioural feedback that can inform or stress-test pathway assumptions. They are therefore included as optional, loosely coupled companions (used when they add explanatory power or improve trigger design) exchanging signals (e.g., adoption rates, supply-chain constraints, learning delays, or financing frictions) with the core decision layer without becoming decision bottlenecks.

Three complementary families are relevant in process heat decarbonisation: System Dynamics (SD), Agent-Based Modelling (ABM), and Stock–Flow Consistent (SFC) monetary accounting. Each offers a way to represent delays, heterogeneity, and financing conditions that shape feasible pathways, while remaining modular enough to "plug in" as light-touch companions to the deep-uncertainty workflow.

5.8.1 System Dynamics (SD)

System Dynamics (SD) represents transition processes as interacting stocks and flows with explicit delays and feedbacks (reinforcing/balancing loops). SD complements RDM and DAPP by providing dynamic insights into transition pathways, enhancing the overall robustness of decision-support under deep uncertainty [359]. For a collection of stocks $\mathbf{x}_t \in \mathbb{R}^n$ (e.g., installed EB/HTHP capacity, trained workforce, regionally mobilised biomass), and net inflow or outflow vectors $f_t^{\text{in}}(\cdot)$, $f_t^{\text{out}}(\cdot)$, a minimal discrete update over time step Δt with first-order delay τ is

$$\mathbf{x}_{t+\Delta t} = \mathbf{x}_t + \Delta t\big(f_t^{\text{in}} - f_t^{\text{out}}\big), \qquad \dot{z}(t) = \frac{z^*(t) - z(t)}{\tau},$$

where $z(t)$ is an adjustment variable (e.g., order backlog or commissioning pipeline) tracking delayed convergence to a desired level $z^*(t)$ [150], [184]. In the New Zealand process heat context, such structures capture procurement and commissioning lags (OEM lead times, contractor capacity), learning/rework cycles after early deployments, and biomass mobilisation with seasonal storage. These feedbacks can amplify or dampen transitions; e.g., ETS-driven biomass orders may raise $z(t)$ faster than logistics can respond, causing short-run shortages before capacity catches up [360], [361].

SD trajectories can supply decision-ready signposts and tipping-point logic for DAPP [9], [50]. For a signpost as a measurable function $s_t = h(\mathbf{x}_t)$ (e.g., available electrified MW, workforce ready-hours, regional biomass drawdown), define a persistence-filtered indicator $\bar{s}_{t,W}$. A trigger fire at

$$\text{Trigger at } t^\star \quad \text{if} \quad \bar{s}_{t,W} \geq s^{\text{thr}} \;\; \forall\, t \in [t^\star - Y + 1, \ldots, t^\star],$$

with a mobilisation lead time L enforcing that the contingent action becomes effective at $t^\star + L$. An SD-informed adaptation tipping point for pathway q is:

$$t^{\text{ATP}}(q) = \inf\{\, t \mid g(\mathbf{x}_t) > g^{\max} \},$$

after which DAPP branches to preserve feasibility [137]. Practically, SD outputs can also inform the time-varying driver vector ω_t used in ensemble stress-tests (e.g., time paths for available capacity or biomass supply), ensuring that exploratory modelling respects real-world frictions [9], [184]. SD adds pathway realism by representing nonlinear feedbacks and delays that ensemble snapshots may miss; it is particularly useful for commissioning bottlenecks, supply-chain lags, and learning curves [184], [361]. Challenges include data-light calibration for New Zealand-specific lags and potential loss of spatial (site) granularity; behaviour-reproduction and structure-oriented validation should be documented [362].

5.8.2 Agent-Based Modelling (ABM)

Agent-Based Modelling (ABM) represents a system as interacting, heterogeneous decision-makers (e.g., firms/sites, suppliers, network operators) whose boundedly rational choices and local interactions generate macro-outcomes over time. This is useful where uptake of low-carbon heat options depends on firm heterogeneity, social influence, supply-chain frictions, and evolving beliefs; features that are hard to capture with aggregate optimisation alone [227], [363].

For the set of industrial sites (agents) S, at decision epoch t, each site $s \in S$ observes a state vector $\boldsymbol{x}_t^s$(e.g., demand, balance-sheet, installed tech), local signals ω_t (e.g., ETS band, delivered-biomass cost, GXP headroom), and peers' recent choices. It chooses an action $a_t^s \in \mathcal{K}_t^s(\omega_t)$ from a discrete option set (status-quo, electrify, biomass, hybrid, defer), where feasibility $\mathcal{K}_t^s(\cdot)$ can encode resource/network limits.

A transparent, stochastic choice rule uses a perceived generalised cost/utility:

$$U_t^s(k) = -\widehat{J}_t^s(k, \omega_t) + \lambda_{\text{pref}}\, \Pi_k^s + \lambda_{\text{peer}} \sum_{s' \in S} W_{ss'} \mathbf{1}\{a_{t-1}^{s'} = k\} \;,$$

with choice probability

$$\Pr(a_t^s = k) = \frac{exp\left(\beta^s U_t^s(k)\right)}{\sum_{k' \in \mathcal{K}_t^s(\omega_t)} exp\left(\beta^s U_t^s(k')\right)}$$

Here $\widehat{J_t^s}(k, \omega_t)$ is the site's perceived discounted cost (or risk) of option k; Π_k^s captures idiosyncratic preferences or strategic fit; W is a row-stochastic influence matrix (sectoral, regional, supply-chain ties); β^s tunes choice noise and discipline; and λ_{pref}, λ_{peer} scale intrinsic and peer effects [364], [365]. Perceptions can adapt with experience (reinforcement learning):

$$\widehat{J_{t+1}^s}(k) = (1 - \rho)\, \widehat{J_t^s}(k) + \rho\, \text{ObsCost}_t^s(k) \;\; , \;\; 0 < \rho \leq 1$$

Adoption is naturally absorbing at the option level. For a binary adoption indicator $A_{s,t} \in \{0, 1\}$ (e.g., electrified or not),

$$A_{s,t+1} = max\left\{A_{s,t},\;\; \mathbf{1}\{a_t^s = \text{electrify}\}\right\}$$

with analogous indicators for other options (biomass, hybrid). Aggregate diffusion and installed capacity then read

$$Y_t = \frac{1}{|S|} \sum_{s \in S} A_{s,t} \;\; , \quad \text{Cap}_t = \sum_{s \in S} A_{s,t}\, \text{cap}_s \;\; ,$$

which can be contrasted with a Bass benchmark $Y_{t+1} = Y_t + (p + q\, Y_t)(1 - Y_t)$ for $Y_t \in [0,1]$, where Y_t is the share adopting a given option, p (innovation) and q (imitation) parameters providing a coarse benchmark [366]. Feasibility can be made endogenous by letting $\mathcal{K}_t^s$ depend on system constraints, for example, if GXP headroom H_t or sustainable biomass B_t (from SD/sector modules) falls below thresholds, certain k are temporarily infeasible or carry penalties in $\widehat{J_t^s}$. This closes a micro–macro loop in which adoption alters constraints that in turn shape future adoption.

ABM outputs supply signposts for adaptive pathways, such as adoption share y_t, backlog of connection requests, or regional biomass draw-down. Triggers can be written as persistence-filtered threshold rules, for example fire trigger at $t^\star$ if

$$\frac{1}{Y} \sum_{\tau=0}^{Y-1} y_{t^\star - \tau} \;\geq\; y^{\text{thr}}$$

(with mobilisation lead time L), or as risk signposts when simulated queue lengths or scarcity indicators exceed bounds, prompting pathway switches (e.g., accelerate biomass where electrification queues persist) [9], [50]. ABM adds behavioural realism (heterogeneous preferences, peer effects, inertia) and exposes path-dependence and lock-ins that static screening can miss. Challenges include calibration of behavioural parameters and influence networks, identifiability, and run-time in large ensembles; good

practice documents behaviour-reproduction tests, parameter sweeps, and rule/ data provenance [363], [364], [367].

In the process heat decision framework, SD and ABM are treated as outer-layer modules rather than core components. Their role is to add behavioural, financial, and institutional realism when data and policy processes permit; they inform signposts and triggers, stress-test proposed pathways, and help prioritise monitoring. Future work could integrate these layers more tightly through hybrid agent-optimisation models or SD-driven feedbacks that update constraints during ensemble runs. Modular separation preserves tractability while ensuring a scalable pathway toward richer socio-technical representations.

5.8.3 Stock–Flow-Consistent (SFC)

Stock–Flow Consistent (SFC) models enforce macroeconomic accounting identities by tracking stocks (e.g., capital assets, financial balances) and flows (e.g., investments, emissions liabilities) in a closed, double-entry framework, ensuring consistency across sectors and time [368]. This is useful in transition planning because decarbonisation pathways hinge not only on techno-economic feasibility but also on financing conditions (credit creation, grants, taxes), with feedbacks from policy (e.g., ETS revenues, investment support) to firms and banks [369].

For the sectors as $\mathcal{Z} = \{F, B, G, H\}$ (firms, banks, government, households), a vector of financial or real stocks at time t is:

$$\mathbf{x}_t^{\mathrm{SFC}} = (K_t^F,\, L_t,\, M_t^H,\, B_t,\, \dots)$$

for firms' productive capital K_t^F, outstanding loans L_t, household deposits M_t^H, and government bills B_t (where dots denote additional instruments such as equity). SFC accounting imposes identities:

$$A\,\boldsymbol{f}_t = 0 \quad (\text{transactions-flow matrix sums}), \quad \Delta\mathbf{x}_t^{SFC} = B\,f_t \quad (\text{stock-flow link}),$$

where $\boldsymbol{f}_t$ collect flows (consumption C_t, investment I_t, government spending G_t, wages W_t, profit Π_t, taxes T_t, interest on loans $i_t^L L_{t-1}$, etc). Where A encodes that every flow is a source for one sector and a use for another, and B maps flow to stock changes (e.g., $L_t = L_{t-1} + \Delta L_t$). A corollary is that aggregate financial balances sum to zero:

$$\sum_{z\in\mathcal{Z}} \mathrm{FB}_{\mathrm{z,\,t}} = 0 \quad ,$$

with $\mathrm{FB}_{\mathrm{z,\,t}}$ the net lending $(+)$, borrowing $(-)$ of sector z [370], [371].

To connect to transition drivers, let firms' operating cash flow incorporate energy/ETS channels:

$$\Pi_t^{\mathrm{ret}} = R_t - \mathrm{OPEX_t}(\omega_t) - P_t^{\mathrm{ETS}}\, E_t(\omega_t) - W_t - i_t^L L_{t-1} \quad ,$$

with exogenous drivers ω_t (e.g., wholesale prices, marginal CO_2, biomass logistics), emissions E_t, and carbon price P_t^{ETS}. A simple investment–finance closure is:

$$I_t^{\mathrm{des}} = \alpha_0 + \alpha_1 q_t + \alpha_2(u_t - u^*) - \alpha_3\, r_t^L + \alpha_4\, S_t, \quad \Delta L_t = \max\{\,0,\, I_t^{\mathrm{des}} - \Pi_t^{\mathrm{ret}} - \Delta E_t\} \quad ,$$

where desired investment depends on profitability q_t, capacity utilisation u_t, loan rate r_t^L, and public support S_t (e.g., grants); ΔE_t captures equity or retained earnings. Government and banking constraints close the loop, e.g.,

$$T_t + P_t^{\mathrm{ETS}} E_t + \Delta B_t = G_t + S_t + i_t^B B_{t-1}, \quad \Delta M_t = \Delta L_t + \Delta B_t - (\text{other instruments}) \quad ,$$

then, ETS revenues and support S_t are financed coherently, and bank expansion accommodates transition credit [368], [370].

For a site $s \in S$, denote investment $I_{s,t}$, retained earnings $\mathrm{Ret}_{s,t}$, grants (e.g., GIDI Fund) $S_{s,t}$, and new credit $\Delta L_{s,t}$. A light financing identity keeps adoption feasible:

$$I_{s,t} = \mathrm{Ret}_{s,t} + S_{s,t} + \Delta L_{s,t}, \quad L_{s,t} = L_{s,t-1} + \Delta L_{s,t}$$

These site-level flows aggregate to I_t and ΔL_t in the SFC shell, allowing portfolio choices (site → region → nation) to respect both technical and monetary feasibility [368], [369].

In the New Zealand process heat context, SFC captures how ETS-driven carbon costs accumulate as liabilities on firm balance sheets, potentially constraining credit for low-carbon investments, while regional biomass flows interact with forestry sector stocks. Feedbacks can amplify transitions; e.g., early electrification success boosts expected profits, accelerating I_t, but GXP constraints or debt build-up may trigger balancing loops (e.g., deferred capex) [370].

SFC trajectories provide signposts for adaptive pathways, such as debt-to-asset ratios or carbon liability stocks. Triggers can be defined as persistence-filtered thresholds, e.g., fire trigger at $t^\star$

$$\frac{1}{Y}\sum_{\tau=0}^{Y-1} f_{t^*-\tau} \geq f^{\mathrm{thr}} \quad ,$$

(with mobilisation lead time L), where f_t is a flow proxy (e.g., investment shortfall). An adaptation tipping point for a pathway q is the first violation of a boundary:

$$t^{\mathrm{ATP}}(q) = \inf\{\; t \mid g(\boldsymbol{K_t}, \boldsymbol{F_t}) > g^{\max}\}\,,$$

prompting switches (e.g., from biomass to electrification if financial stocks signal stranded coal risks) [9].

SFC adds macroeconomic realism by ensuring financial-physical consistency and exposing systemic risks (e.g., transition-induced recessions) that micro-focused models miss; it is especially useful for policy feedbacks like ETS revenue recycling or green financing in NZ contexts. Challenges include data requirements for balance-sheet calibration and potential over-aggregation of site heterogeneity; validation should include stock-flow audits and sensitivity to behavioural closures [368], [370]. In the context of modelling decarbonising process heat SFC could complement RDM and DAPP by providing consistent macro insights into transition pathways, enhancing the overall robustness of decision-support under deep uncertainty.

Figure 5.6 recapitulates the role of socio-technical modules as optional, outer-layer companions to the deterministic–exploratory core. System Dynamics (SD), Agent-Based Modelling (ABM), and Stock–Flow-Consistent (SFC) accounts generate dynamic, behavioural, and financing signals (e.g., capacity paths, adoption, queues, credit, liability flows) that sharpen signposts and trigger thresholds in DAPP. In return, the core passes back candidate policies and trigger rules for stress-testing. This placement preserves tractability in the core while enabling escalation to richer socio-technical realism when warranted by data or governance needs.

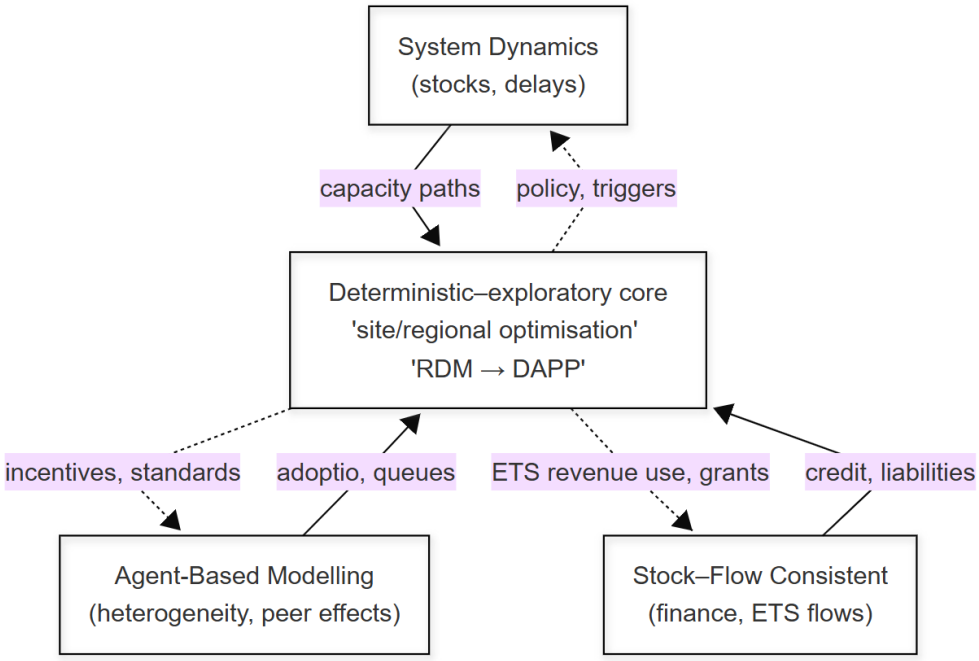


Figure 5.6 Socio-technical modules (SD, ABM, SFC) interfacing with the decision framework

5.9 Implications for framework design

Industrial process heat decisions are high-stakes, long-lived, and difficult to reverse. Assets persist for decades while key conditions (including ETS credibility and price bands, delivered-biomass costs and sustainability, GXP hosting capacity and reinforcement timing, technology learning, and demand trajectories) remain uncertain and sometimes contested. In this context, "informed" decisions cannot rely on single forecasts or narrow sensitivity checks. They require a decision framework that:

- establishes engineering feasibility with traceable accounting,
- exposes how strategies perform across wide ranges of plausible futures,
- makes failure conditions explicit and monitorable, and
- supports adaptive, staged implementation consistent with procurement lead times and governance realities.

The foregoing review implies the building blocks for such a framework and clarifies their roles. Taken together, these approaches emphasise a framework that is modular by construction, exploratory in posture, and auditable in practice. As illustrated in Figure 5.7, the deterministic optimisation remains the authoritative feasibility engine, enforcing engineering and policy constraints, emitting cost and emissions ledgers, and exposing binding conditions (e.g., GXP headroom, sustainable-biomass envelopes). Where credible distributions exist, a probabilistic–stochastic inner loop quantifies near-term operational risk (e.g., Monte Carlo and L-shaped/Benders recourse) and reports quantiles, CVaR, and violation shares. Under deep and structural uncertainty, ensemble-based robustness becomes the organising layer (RDM/MORDM), with evidence reported as satisficing coverage, regret percentiles, and reliability shares rather than point optima. Scenario discovery compresses ensemble results into sparse, auditable vulnerability statements that feed Dynamic Adaptive Policy Pathways (DAPP) as signposts and trigger thresholds. MCDA provides a transparent evaluation layer that keeps robustness metrics decision-visible, supports non-compensatory checks where standards are non-negotiable, and can be lifted into budget-constrained portfolio selection to preserve system couplings. Value-of-Information (EVPI/EVPPI/EVSI) focuses data collection and model refinement where they change decisions. Finally, socio-technical companions, including System Dynamics (SD), Agent-Based Modelling (ABM), and Stock–Flow-Consistent (SFC) shells, operate as optional, loosely coupled modules that supply dynamic, behavioural, and financing signals to sharpen triggers and stress-test pathways without overburdening the core. Figure 5.7 synthesises this decision-centred integration and coupling.

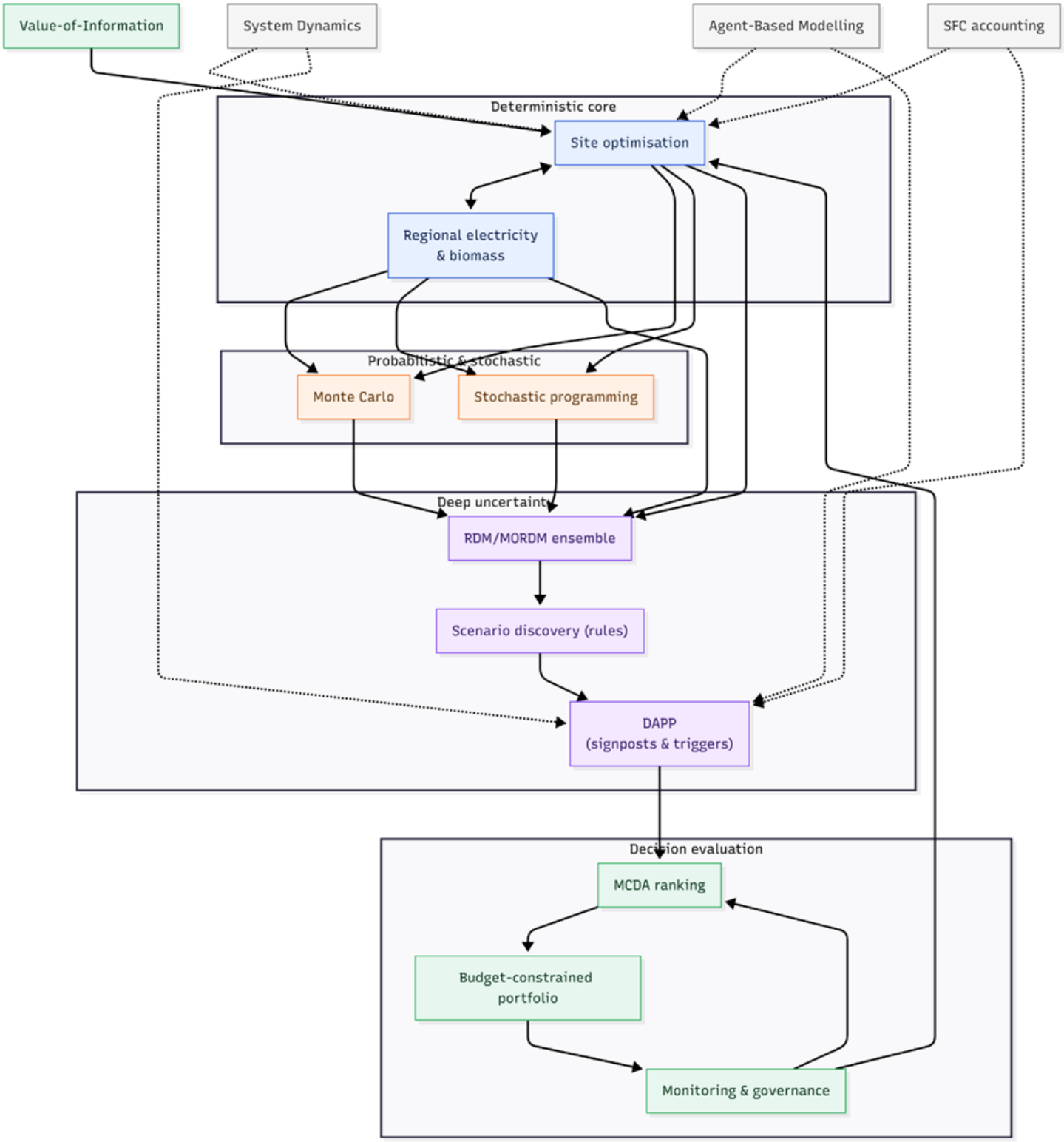


Figure 5.7 Integrated, modular decision framework for process heat decarbonisation

Recent studies on low-regret strategies reinforce this integration, highlighting robustness in volatile policy contexts for industrial transitions [372]. Operationally, this synthesis yields concrete design requirements:

- **Interfaces and artefacts:** Each run should emit a standard artefact pack; KPIs, constraint and dual logs, robustness dashboards (satisficing, regret, reliability), MCDA-ready performance matrices; scenario-discovery rule sheets with coverage, precision, support and validation stats, and provenance manifests (data ranges, structural toggles, solver versions, seeds, checks).
- **Sampling and hierarchy:** Use space-filling designs for ensembles; evaluate hierarchically (site → region → national); re-check boundary or contentious cases at full chronology; and prune behaviourally similar futures to control run time while preserving coverage, essential in New Zealand context where site outcomes couple tightly to regional constraints.
- **Adaptivity:** Translate discovered thresholds into monitorable signposts with persistence windows and mobilisation lead times; pair these with pre-agreed contingent actions in DAPP pathway maps and scorecards.
- **Evaluation and portfolios:** Keep robustness columns (regret, satisficing, reliability) inside the decision matrix; apply non-compensatory vetoes where minimum standards (e.g., reliability, sustainability) must not be traded away; embed rankings in budget-constrained portfolio selection that respects regional caps.
- **Governance and legitimacy:** Support co-production of criteria, weights, and trigger calibration; expose value-function anchors and weight sweeps; maintain acceptance tests and traceability so figures and tables resolve to inputs and code states.
- **Extensibility:** Keep SD/ABM/SFC light-touch, escalating for bottlenecks (e.g., adoption inertia), without decision bottlenecks, facilitating application beyond process heat.

By design, the framework concentrates effort where it moves decisions; binding constraints in the core, high-leverage uncertainties in ensembles, and implementable triggers in pathways. These implications define practical requirements for data, modules, and interfaces, informing credible application to New Zealand industrial process heat and previewing requirements and gaps. This thesis terms the resulting framework Decision-Centred Modelling (DCM); Chapter 6 specifies its architecture.

Chapter 6
Architecture of the Framework: Requirements, Interfaces, and Artefacts

6.1 Introduction

Industrial process heat is among the most complex and least tractable domains of contemporary energy systems, characterised by diverse technologies, site-specific configurations, and tight dependencies on regional fuel supply, infrastructure, and policy settings. Unlike the electricity sector, where balancing and market rules are centralised, process heat transitions arise from distributed, path-dependent choices made by individual firms under uncertainty, with decisions at specific facilities propagating through regional electricity and biomass systems (and any other alternatives) whilst regional constraints feed back into the credibility and cost of site-level options. Decarbonising this sector therefore demands an analytical architecture that links local operational realism with system-level constraints and reveals how strategies perform as external conditions shift, all while remaining tractable, transparent, and decision-relevant. The approach adopted here rests on a simple premise: carry complexity where it creates decision value and abstract it where it would only obscure traceability.

The framework proceeds from modular decomposition. Site-level models represent processes, utilities, and operational constraints with the temporal fidelity required to establish feasibility and to quantify cost and emissions. Regional modules aggregate across sites and impose shared limits (fuel availability, policy caps, network and capacity constraints), whilst returning scarcity signals that reflect the marginal value of those limits. Between these tiers a deliberately narrow set of interfaces (the thin waist) conveys only what must be shared; demand and profile artefacts, system signals and policy parameters, and a compact evaluation outlook. By contracting information in this way, the architecture preserves freedom for method choice within modules while guaranteeing coherent exchange across them.

Coordination is deliberately light. Rather than collapse all decisions into a monolithic programme, the architecture uses a small number of Gauss–Seidel-style passes in which quantities and shadow prices are exchanged between tiers until consistency is achieved to tolerance. This yields auditable solutions with transparent constraint placement: sites solve their feasibility problems against externally supplied signals; the region solves for shared constraints given aggregated site behaviours; and dual variables

provide the economic language through which iterative coherence is restored. The result is a structure that avoids over-coupling while retaining the essential physics and economics of sectoral interaction.

Uncertainty is addressed by layering exploratory analysis around the deterministic core, not by entangling it within. Decision epochs reflect the staged nature of industrial investment; ensembles of plausible futures articulate uncertainty in fuel costs, policy trajectories, demand, and network headroom; and robustness indicators, such as regret, reliability, and satisficing, evaluate strategies across these futures. This separation of concerns keeps the core solvable, testable, and reusable, while the exploratory layer provides the breadth required for adaptive planning. In the New Zealand setting, where emissions pricing credibility, biomass deliverability, and grid constraints at GXPs are material and evolving, such a posture is essential for credible guidance.

The claims advanced in this chapter are conceptual and practical: that a thin-waist modular architecture can deliver scalability without loss of coherence, auditability without forfeiting temporal realism, and decision-useful robustness without inflating the core problem. These qualities are not stylistic preferences but necessary conditions for models intended to inform real industrial conversions under policy and infrastructure constraints.

Key terms are used sparingly and in their operational sense; formal definitions are provided in Appendix 6-A, and mathematical detail follows in the next chapter. What follows articulates the architecture itself: the constituent tiers, their interactions, and the information contracts that allow site-level realism to co-exist with regional limits in a manner that remains tractable and fit for decision support.

6.2 Architectural overview

The architecture set out in this chapter provides a disciplined way to link local operational realism with system-level constraints for New Zealand's industrial process-heat transition. Modular decomposition is necessary because salient frictions, such as volatile ETS price bands, uncertain biomass deliverability, and spatially heterogeneous GXP headroom, operate at different scales yet interact strongly. The design goal is to exchange only the information that must cross scales, so that site models remain faithful to plant realities while the regional layer enforces shared limits and returns interpretable signals.

Tiers and their responsibilities

The architecture has three operative tiers, with an optional national wrapper for validation.

Site tier. Each site module represents processes and utilities with the temporal fidelity required to establish feasibility at an hourly chronology. It owns constraints that are intrinsically local: capacity and ramping limits, on-site balances, temperature or service requirements, fuel switching, and plant-level cost and emissions accounting. Its output is a clean disclosure of what the site demands from shared systems under stated signals.

GXP/sub-regional tier. Where network connection or fuel logistics are organised below the regional level, an intermediary hub aggregates site outputs by GXP or logistics catchment. It checks local headroom or handling limits and relays concise signals up and down without duplicating regional optimisation.

Regional tier. The regional module owns constraints shared across sites: biomass availability and sustainability limits, policy caps and accounting rules, and electricity network and capacity limits. It aggregates site disclosures, enforces shared constraints, and returns interpretable scarcity signals (shadow prices) that reflect the marginal value of constrained resources and infrastructure.

Optional national validation layer. A light wrapper compares or cross-checks regional outcomes against national references or alternative back-ends; it does not impose additional constraints.

Constraint ownership is thus explicit; plant realism and immediate feasibility reside at the site; shared resource and policy limits reside at the region; locality-specific checks may be mediated at GXP/sub-regional level; national wrappers serve validation rather than control.

The thin-waist interfaces

Information moves across tiers via three versioned artefacts with narrow, stable schemas. The interface is intentionally small: only essential fields cross tiers; schemas are versioned and change rarely; modules remain decoupled internally provided they conform exactly to the published artefact specifications (fields, units, identifiers, chronology, and schema version). Each exchange is released only after passing automated acceptance checks (schema conformance, identifier integrity, unit/chronology validation, and hash/version verification), so every transfer is recorded and auditable.

- **DemandPack** transmits site-produced demand and configuration (hourly or representative-day chronology with weights) upward for aggregation and shared-constraint evaluation.

- **SignalsPack** transmits downward the regional context that defines the site problem: time-indexed prices, headroom/envelopes, and emissions limits provided ex ante (before a site run). After regional evaluation, optional scarcity indicators (e.g., shadow values) may be returned ex post to inform the next iteration. Sites then solve their optimisation conditional on the contents of the SignalsPack they receive.
- **OutlookPack** carries upward the results needed for evaluation and reporting: feasibility flags, ledgers, KPIs, and concise summaries of binding constraints and scarcity indicators, including fields used for robustness panels when ensembles are run.

Figure 6.1 situates the tiers and the interfaces they exchange.

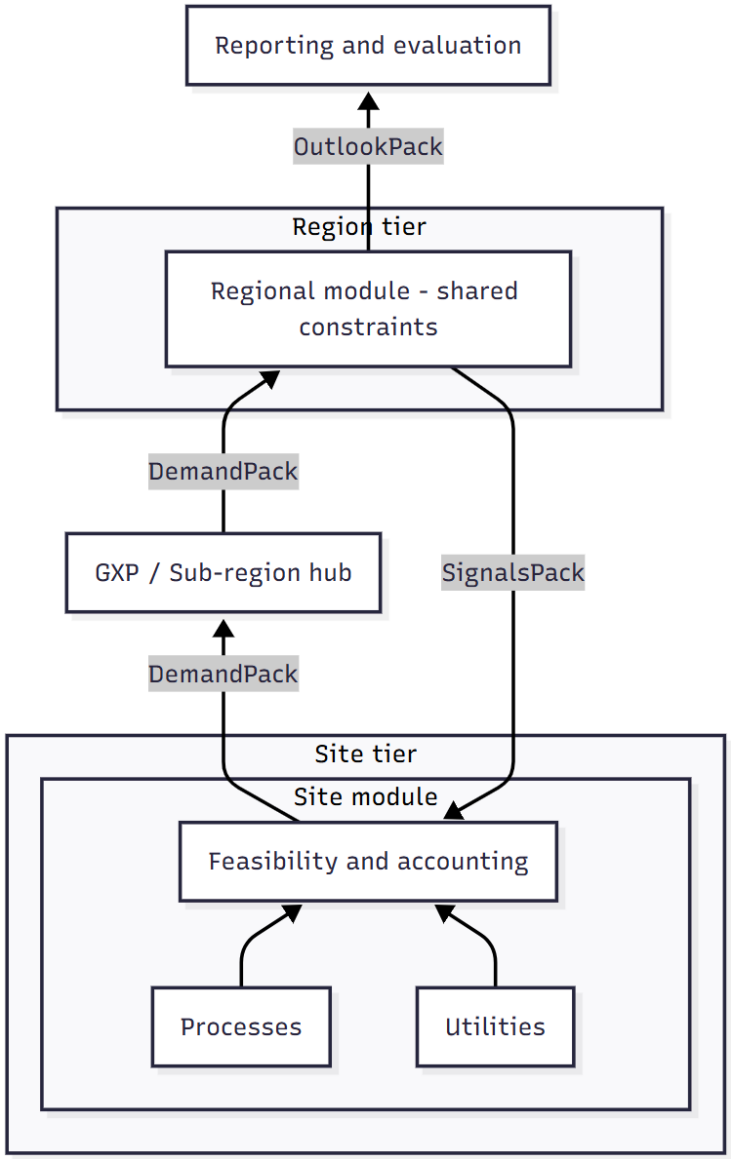


Figure 6.1 Multi-scale architecture and thin-waist information flow

Coordination loop (light exchange to tolerance)

Rather than collapse all decisions into a monolith, the tiers coordinate through a short price–quantity exchange. An initial SignalsPack provides policy and network context to sites; each site solves feasibility and publishes a DemandPack. The region aggregates these disclosures, enforces shared constraints, and issues an updated SignalsPack that includes scarcity signals where constraints bind. Sites respond, potentially adjusting operating modes or technology selections, and republish DemandPack. The loop repeats until pre-declared tolerances are met, for example, negligible change in aggregate demands or regional duals (the shadow values associated with regional/shared constraints such as GXP headroom, biomass envelopes, or emissions caps; they indicate the marginal cost of relaxing the constraint by one unit), and no change in binding status (whether a constraint is tight within tolerance and carries a materially non-zero dual), delivering coherence and interpretability while keeping modules independently verifiable between passes.

Temporal and spatial scaffold

Operational realism at the site tier is maintained through an hourly chronology within baseline years. Decision epochs (e.g., 2025, 2030, 2040, 2050) provide the rhythm for adaptation as signals and observed conditions evolve. Where computation requires it, hours may be compressed into representative periods with explicit diagnostics on energy, peak, ramp, and storage fidelity. Spatial indices progress from site → GXP/sub-region → region, aligned with the ownership of constraints described above.

Adapters and governance (pointers)

External back-ends integrate via adapters that read and write the same artefacts without altering their schemas (e.g., delegating network checks to a PyPSA-style engine or running planning comparisons with a TIMES-style formulation). Artefact exchanges are guarded by manifests, acceptance gates, and a run ledger to ensure provenance and reproducibility

6.3 System boundary and scope definition

This section defines the operating envelope within which industrial process-heat decisions are evaluated. It fixes where constraints bind (spatial tiers) and when choices are assessed (operational chronology and decision epochs), and it specifies the decision domains that give those scales meaning. Two complementary perspectives are maintained: a top-down view that frames decision relevance

(technology change, fuel substitution, infrastructure coordination, emissions management, policy settings) and a bottom-up view that anchors feasibility in engineering and logistics (energy and mass balances, unit capabilities, GXP headroom, biomass envelopes). Together, these ensure that explored strategies are both decision-useful and physically realisable.

6.3.1 Spatial scope and carriers

The analysis uses a nested hierarchy (site → sub-regional → regional → national), so that constraints are enforced where they are materially relevant and signals are generated at the appropriate level.

At the site boundary, the control volume comprises industrial processes and on-site utilities such as boilers, heat pumps, and thermal storage. Heat and electricity balances must close at the chosen chronology; temperature or service requirements and equipment limits (capacity, ramping, availability, maintenance) are enforced locally. External networks and upstream fuel production are outside the boundary. Cross-boundary inflows are purchased electricity at the connected Grid Exit Point (substation) and delivered fuels such as biomass; where modelled, exports are treated explicitly. The site tier is intentionally general and can also represent other end-use domains (residential areas or transport depots) provided the same discipline of control volumes, balances, and disclosures is applied.

The GXP–sub-regional boundary, used where material, groups proximate sites for connection access and local logistics checks. Inside this boundary sit the connection queue, transformer and feeder headroom, and any explicitly modelled biomass logistics envelopes covering harvesting, processing, storage, and seasonal transport. Plant-internal dispatch and region-wide policy caps are outside. Flows crossing the boundary include aggregated site net loads and fuel draws upward, and concise access indications downward such as headroom signals or curtailment flags. This tier performs checks and relays information; it does not duplicate regional optimisation.

At the regional boundary, the control volume contains shared resource pools and policy accounts. Electricity adequacy and network capacity limits at the chosen aggregation, sustainable biomass availability with its sustainability conditions, and regional emissions budgets with their accounting rules bind here. Site-internal operations are outside. Cross-boundary exchanges comprise aggregated demands and resource claims from below, and policy or scarcity signals returned downward where limits bind. The regional scope is general and may include other carriers where they are materially shared, including natural gas deliverability, geothermal resource availability, hydrogen generation with licensing constraints, and regional heat networks, provided each carrier’s control volume and binding limits are defined consistently.

The national boundary functions as a policy and validation wrapper rather than an additional optimisation layer. It holds cross-regional conventions such as emissions-trading price bands, eligibility rules, and national targets, and provides a common reference for reporting and comparison with alternative national models. It does not alter the regional solves but supplies the accounting baseline against which regional emissions and costs are tallied.

The current focus is process heat coupled with electricity and biomass, reflecting technical plausibility and data availability. The scaffold remains extensible across both carriers and demand sectors under the same boundary discipline. These relationships are demonstrated in Figure 6.2.

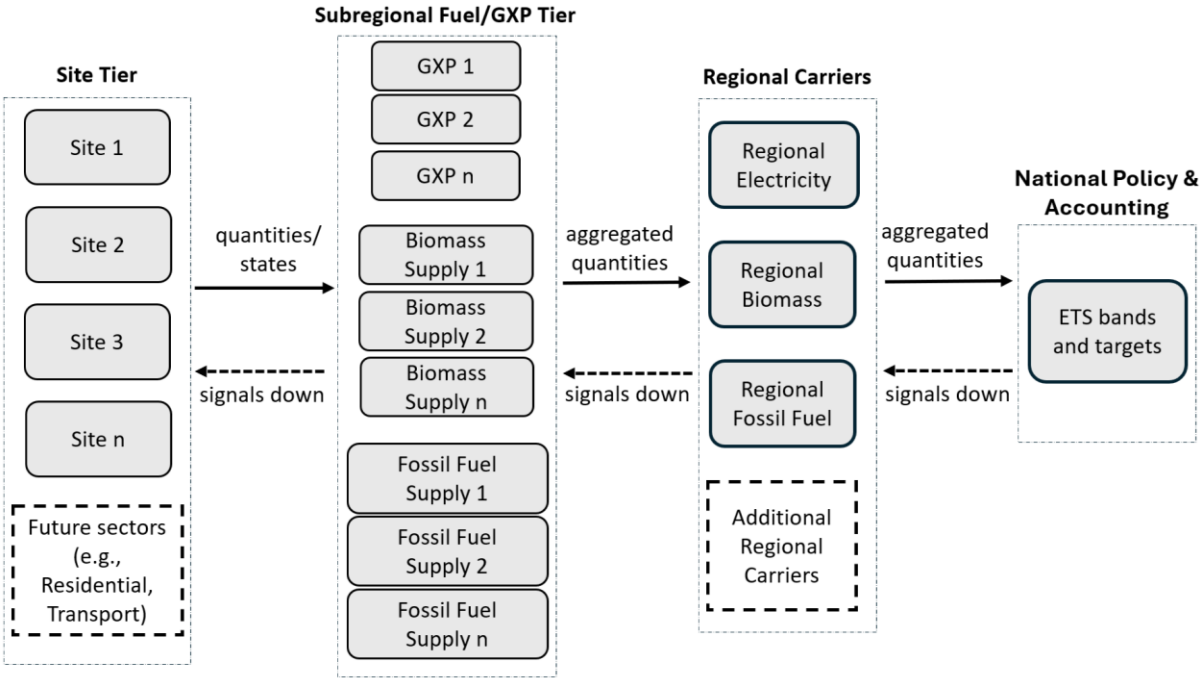


Figure 6.2 Nested spatial tiers with exemplar constraints and exchanged signals

6.3.2 Temporal scaffold and decision epochs

Time is treated on two layers within the framework, which distinguishes the operational chronology within a year from the multi-epoch planning horizon across years, and it specifies the state variables that carry system memory between epochs. The aim is to keep feasibility checks transparent at the operational scale while enabling credible sequencing of investments and policies over time.

At the operational level, the model evaluates site feasibility on an hourly chronology for the baseline year, or on a defensible set of representative periods whose weights sum to annual hours. Representative sets must preserve seasonality, peak conditions, typical weekday–weekend patterns, ramping behaviour, and storage cycling. Mapping rules and weights are documented so that annual

energy, peak demand, and utilisation factors can be reproduced within declared tolerances. Where representative periods are used, boundary conditions for stateful units are specified explicitly to avoid artefacts (cyclic end-states for storage and clear treatment of partial segments), so that feasibility is neither spuriously relaxed nor made infeasible by construction (i.e., infeasibilities that arise purely from the compression/mapping or from mis-specified boundary conditions rather than from the underlying physics or policy constraints). Time conventions (local clock, handling of daylight-saving transitions where relevant) are fixed once and applied consistently across sites.

Across years, decisions are organised into discrete epochs, for example 2025, 2030, 2040, and 2050, as shown in Figure 6.3. Within an epoch, the operational chronology is evaluated under the signals in force for that epoch. Between epochs, the framework updates state variables that record the system's memory: installed capacities by technology and their vintage or remaining life; cumulative emissions against the applicable accounting rules; committed reinforcements or contracts; and any deferred retirements. Lead times (the delay between decision and availability in service) and lumpy investments (indivisible, discrete capacity blocks, e.g., a 30 MW boiler often subject to minimum size thresholds and fixed start-up costs, hence not addable in arbitrary fractions) are handled at the epoch boundary, so that option value and path dependence are represented without inflating the operational solve. Where adaptive planning is examined, signposts and triggers are defined ex ante to govern admissible recourse between epochs, subject to procurement and consenting constraints.

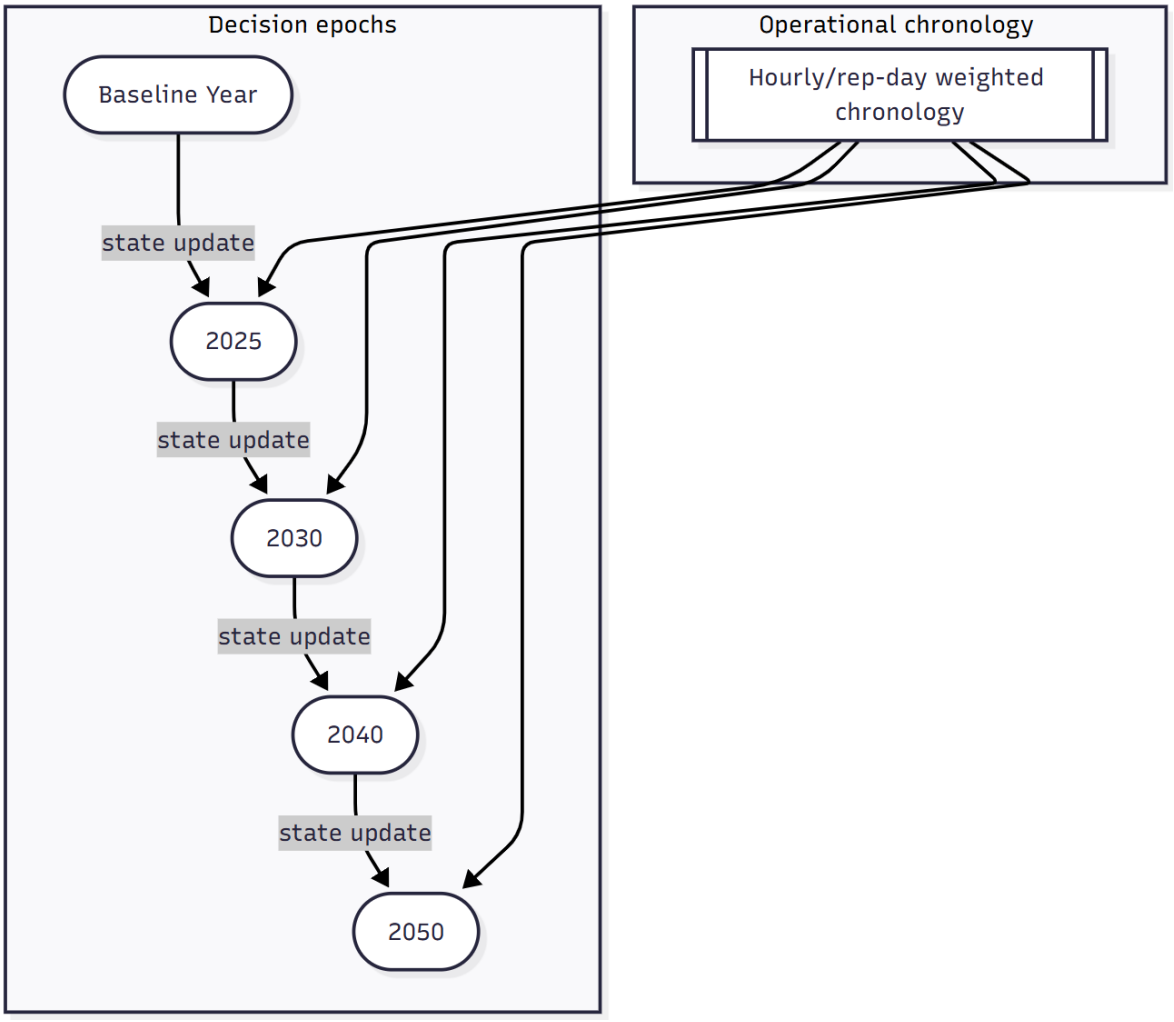


Figure 6.3 Temporal scaffold and decision epochs

Uncertainty is treated around, rather than inside, the operational core. Exogenous draws such as emissions-price bands, tariff schedules, biomass trajectories, and network headroom enter the epoch as part of the external signals; endogenous scarcity signals returned from shared constraints are recorded with the operational results. Ensembles of such conditions can be run across epochs, with the reporting layer collating performance and robustness indicators. Methods for selection of representative periods, ensemble design, and robustness reporting are specified elsewhere; here the concern is that the chronology and state transitions be defined in a manner that is auditable and stable.

Compression is optional and governed by diagnostics. If hours are compressed, the selection process, weights, and acceptance criteria are declared, and post-hoc tests confirm fidelity for energy, peaks, ramps, and storage. These rules are applied uniformly to all sites within an epoch so that aggregation to the regional layer remains coherent. The separation of intra-year operations from inter-epoch evolution provides a clear temporal scaffold: feasibility is demonstrated on a chronology that preserves

the salient operational features, while investments, retirements, and policy shifts are sequenced at a handful of decision points with explicit state carry-over. The following maps concrete decisions to this scaffold by specifying who decides what at each tier, which quantities are permitted to cross temporal and spatial boundaries, and how those exchanges are recorded for subsequent evaluation.

6.3.3 Mapping decisions to boundaries and information exchange

The analytical boundary of a DCM application is set by the decision, not by the physical extent of any one system. The decision-relevant boundary is the smallest analytical scope within which the consequences capable of changing the comparison of the available alternatives are visible. The principle is selective rather than expansive: it reaches across scales to include a consequence that bears on the comparison, such as grid reinforcement triggered by a site electrification choice, while excluding elements at the same or larger physical scale that do not. Because the relevance of a consequence cannot always be known before it is examined, the boundary is recorded rather than assumed. Each candidate consequence is classed as included (modelled and carried into the comparison), proxied (represented at reduced fidelity), excluded (outside the boundary, with a stated reason and a route by which it could enter later) or unresolved (its relevance cannot yet be judged). The record is revised as the analysis's own diagnostics show which consequences move the comparison. Section 7.4.2 gives this boundary register for the proof of concept.

Decisions are located at their proper tier, constraints bind once at the tier that owns them, and only the minimum necessary information crosses each boundary. The purpose is to make the decision contract explicit so that subsequent formulation can remain concise; feasibility and dispatch are addressed where they belong; shared limits are enforced without duplication; exchanges across tiers are deliberately sparse.

At the site tier, decision rights cover technology adoption and sizing within the enumerated options, operational scheduling of processes and utilities at the chosen chronology, and the operation of storage subject to local capability limits. These decisions are taken conditional on external signals for the epoch. The site's obligation to higher tiers is a disclosure of time-indexed net demands on shared systems, together with on-site fuel use and emissions computed under agreed accounting conventions. That disclosure is issued as a versioned DemandPack and is the site's only routine upward message; internal states and heuristics remain private.

The GXP and sub-regional tier, where employed, does not make optimisation decisions. Its role is to aggregate site disclosures, test them against local connection headroom and explicitly modelled fuel logistics, and return concise access indications to affected sites. These may include indicative

connection timelines, headroom signals, or curtailment flags. Plant-internal choices are not revisited here, and region-wide policy constraints are not duplicated. The tier acts as a thin mediation layer to prevent infeasible local congestion from propagating to the regional solve.

At the regional tier, decision rights cover the enforcement of shared constraints across sites. The region aggregates site disclosures, applies adequacy and capacity limits, imposes sustainable biomass availability and any environmental conditions, and accounts for regional emissions budgets. Where limits bind, the region computes scarcity or shadow prices that quantify the marginal value of the constrained resource or capacity. The region's outward communications are twofold: a SignalsPack issued downward, containing the policy and network context for the epoch and, after a pass, any scarcity signals; and an OutlookPack issued to the reporting layer, containing key performance indicators and, where ensembles are run, the carriers required for robustness assessment. No site-internal variables are imported, and no plant-level constraints are altered at this tier.

National conventions set the accounting baseline rather than additional optimisation decisions. Emissions-trading price bands, eligibility rules, and national targets are fixed here and referenced by the region when computing compliance and by the reporting layer when collating results. The wrapper may also provide a validation point against alternative national models, but it does not inject constraints into the regional pass.

The information that crosses boundaries is therefore deliberately narrow. Upward exchanges comprise quantities and states strictly necessary for aggregation and enforcement, including demands and profiles, fuel draws, emissions summaries, and any capacity requests or commitments that carry between epochs. Downward exchanges comprise signals and envelopes strictly necessary for local feasibility and choice: tariff schedules and emissions-price bands, indicative connection limits, resource envelopes, and scarcity prices where limits bind. No other internal details are exposed. This separation maintains sufficiency and auditability while preserving method freedom within modules.

Exchanges are objective-aware and configuration-driven. Each artefact has a versioned core schema with optional, namespaced facets that are enabled only when required by the declared decision objective. A run manifest selects the objective and activates the corresponding profile, so modules read and write only the facets they declare. This keeps the waist thin engaging only the data necessary for the question, while preserving comparability through a baseline indicator set that is always present in OutlookPack.

6.4 Functional and analytical requirements

Building on the spatial tiers, temporal chronology, and information exchanges defined earlier, this section states normative, testable requirements for a tractable and decision-relevant evaluation of industrial process heat. Two complementary lenses guide the specification: a top-down view that identifies the decisions to be supported (technology change, fuel substitution, infrastructure coordination, emissions management, and policy settings) and a bottom-up view that anchors feasibility in physics and logistics (energy and mass balances, unit capabilities, regional envelopes, and network headroom), so that candidate strategies remain operationally credible rather than merely descriptive.

To keep the design coherent yet comprehensive, the requirements are organised into four clusters, each decomposing into the seven atomic requirements (see Table 6.1):

Cluster A — Feasibility, coordination, and scale: Deterministic feasibility and accounting at site and regional levels (R1); site–region coherence through quantity–signal exchange with scope for selective co-optimisation where couplings are tight (R2); tractability and scalability from desktop screening to parallel ensembles (R3).

Cluster B — Exploratory robustness and adaptive pathways: Designed ensembles and discovery to expose performance ranges and failure modes (R4); sequencing of long-lived choices with signposts and triggers for recourse between decision epochs (R5).

Cluster C — Provenance and auditability: Versioned manifests, a run ledger, and pre-/post-solve acceptance checks for inputs and outputs to support independent replication (R6).

Cluster D — Evaluation readiness and comparability: A stable panel of metrics (cost, emissions, adequacy, utilisation, and robustness measures such as regret, reliability, and satisficing) enabling consistent comparison across sites, epochs, and ensembles (R7).

Table 6-1 Requirements, rationale, and indicative implementations

Requirement	Rationale (decision relevance)	Implementation implication (generic handle)	Acceptance (evidence)
R1: Feasibility & accounting	Physically and policy-consistent results	Enforce balances/limits; explicit emissions ledger	Feasibility certificate; constraint log; emissions ledger reconciles to aggregates
R2: Site–region coherence	Local choices must respect shared limits	Quantity–signal exchange; optional targeted co-optimisation	Convergence to named tolerances; duals reported and interpretable
R3: Tractability & scale	Practical execution and extensibility	Compression; parallelism; caching; declared tolerances	Run ledger records runtime/optimality/accuracy; desktop single-scenario solvable
R4: Uncertainty ensembles	Performance under many plausible futures	Scenario register; designed sampling; seed control	Reproducible ensemble; coverage diagnostics pass; results trace to seeds
R5: Adaptive sequencing	Path dependence and option value	Epoch states and lead times; ex-ante signposts/triggers	State continuity checks pass; trigger logic well-posed; admissible recourse only
R6: Provenance & audit	Reproducibility and trust	Manifests; run ledger; schema validation	Independent rerun reproduces results (within declared numerical tolerances)
R7: Metrics	Comparable evaluation and robustness	Stable KPI panel; regret/reliability/satisficing	KPIs present for all runs; aggregation rules/units explicit

6.4.1 Deterministic feasibility, coordination, and scale (R1–R3)

The framework shall establish site-level feasibility on the chosen operational chronology (hourly or approved representative periods). Heat and electricity balances must close; temperature or service requirements must be met; equipment limits on capacity, ramping, start-up and shut-down, storage continuity (state of charge and inter-block linking), availability and maintenance derating must be respected. Policy compliance is enforced through explicit emissions accounting (scope definitions, factors, caps or budgets), technology-eligibility rules, and any prohibition on new fossil capacity.

Each run produces (i) a feasibility certificate confirming that balances close within named tolerances, (ii) an emissions ledger by source and scope that reconciles to aggregates, and (iii) a constraint log that identifies binding and near-binding limits.

Constraints bind where they are materially owned. Site modules enforce process and utility physics and produce time-indexed quantities (dispatch, fuel use, emissions, capacity requests). Sub-regional and network checks enforce access and adequacy envelopes, for example Grid Exit Point headroom and feeder limits; logistics modules bound deliverable biomass via cost–quantity envelopes and seasonal availability. Regional carrier modules aggregate quantities and apply shared limits on adequacy, sustainable biomass and regional emissions budgets. Information flows upward as quantities and states; information flows downward as envelopes and interpretable scarcity signals (for example indicative headroom, biomass scarcity rents, emissions shadow values). Dual variables associated with shared constraints are reported as signals without assuming any specific market design.

Consistency between local decisions and shared limits is achieved through a light price–quantity exchange. Given current signals, sites re-optimise; given updated site outputs, sub-regional or regional modules recompute envelopes and scarcity indicators. Convergence is declared when named tolerances on balance closure and on successive scarcity signals are met, or when a recorded maximum iteration count is reached. Tolerances and limits are defined and recorded for each run in the ledger. Where couplings are tight (for example a small set of sites behind a constrained Grid Exit Point), selective co-optimisation is confined to the affected subproblem while other modules remain unchanged; the scope and constraints of such cases are identified in the ledger. This preserves modular transparency, admits heterogeneous model detail, and avoids unnecessary joint solves.

Emissions, energy, and cost accounting are explicit and keyed for audit:

Energy ledger: input fuels, conversion and storage losses, and delivered heat and electricity.

Emissions ledger: factors by carrier and scope, annual and cumulative totals, and cap usage.

Cost ledger: annuitised capital expenditure, fixed and variable operating expenditure, fuels, network charges, and emissions trading charges.

Computational burden is controlled without compromising credibility. Temporal compression (weighted representative periods) is permitted where justified, with explicit boundary conditions for stateful assets and documented reconstruction error for annual energy, peaks, ramps, and storage cycling; spatial aggregation (for example grouping very small sites or feeders) is admissible only where it does not distort binding constraints. Problem size is contained by preferring linear formulations, reserving integer variables for genuinely discrete build or retire decisions, caching invariant matrices, reusing warm starts across iterations and scenarios, and executing ensembles in parallel. A single-scenario deterministic run shall remain solvable on a modern desktop.

The run ledger records solver settings, runtime, and optimality and feasibility tolerances; compression diagnostics pass stated thresholds; any spatial aggregation is justified in the manifest. The feasibility stance shall be declared: either prohibited (hard feasibility) or permitted with a stated penalty. The policy is recorded in the configuration manifest and reflected in the feasibility certificate.

Each run reports: feasibility status; violated or near-binding constraints with magnitudes; utilisation of headroom and envelopes by tier; summaries of duals or indicative prices for shared constraints; and basic performance indicators (energy served, unmet heat where admissible, utilisation factors). Where infeasibilities occur, the minimal conflicting constraint set is reported to guide correction of data or assumptions rather than masked by slack variables.

Together these provisions deliver feasibility and accounting discipline (R1), ensure site–region coherence through transparent exchange and auditable convergence (R2), and preserve tractability and scale through prudent formulation, compression, and decomposition (R3).

6.4.2 Exploratory robustness and adaptive pathways (R4–R5)

Uncertainty in prices, demand, technology performance, infrastructure timing, and policy warrants an explicit exploratory layer. The aim is not to predict a single future, but to evaluate how candidate strategies behave across a structured ensemble and to allow recourse when conditions diverge from expectations.

A scenario registry lists uncertain drivers with ranges, distributions, units, provenance, and epoch applicability. Ensembles are generated using defensible design-of-experiments methods (for example Latin hypercube or Sobol) with seed control for reproducibility. Each scenario bundles a complete, internally consistent set of exogenous inputs at operational and epoch scales. Where temporal compression is used, the same representative set and weights are applied across the ensemble to avoid confounding sampling with chronology. Exogenous draws are carried in SignalsPack; the core returns feasibility status, key performance indicators (total cost, emissions, adequacy and utilisation), and dual/constraint diagnostics; OutlookPack collates outcomes and robustness carriers.

For each scenario, the deterministic core returns feasibility status, key performance indicators (total cost, emissions, adequacy/utilisation), and dual/constraint diagnostics. Robustness is then assessed using a small, standardised panel:

- **Regret:** difference between the outcome of a chosen strategy and the best feasible strategy in the same scenario.

- **Reliability:** share of scenarios in which pre-declared satisficing thresholds are met.
- **Satisficing:** whether all thresholds for the KPI panel are simultaneously satisfied.

Simple discovery tools (for example PRIM or CART) label vulnerability regions (combinations of drivers that systematically yield infeasibility or poor performance) guiding strategy refinement and data-collection priorities. Discovery is diagnostic; it does not alter the core.

Long-lived choices are evaluated as pathways across decision epochs. Signposts track observable indicators (such as realised ETS band, achieved GXP headroom, biomass delivered cost); triggers define thresholds at which recourse becomes admissible (for example defer electrification, add storage, pivot to biomass). Recourse respects lead times, stock dynamics, and eligibility rules. Pathways are assessed by replaying the ensemble with trigger logic active, so sequencing quality (not only static portfolios) is tested. Outputs include pathway maps, trigger hit-rates, and distributions of cumulative cost and emissions with adaptations applied.

The exploratory layer remains loosely coupled; it perturbs contextual inputs, orchestrates calls to the core, and aggregates outputs without modifying feasibility logic. When a run fails, failure labels and the minimal conflicting constraint set are retained rather than relaxed, preserving diagnostic value. For adaptive analyses, epoch states (installed stocks, remaining life, cumulative emissions) are carried forward exactly as returned by the core.

Every scenario run is recorded in the run ledger with a unique identifier, seed, parameter realisation, software versions, and artefact hashes. Acceptance checks verify that scenario bundles are complete and within bounds, that ensemble coverage is adequate (simple space-filling diagnostics), and that trigger definitions are well-posed (no contradictions; feasible lead times). Parallel execution is applied at the scenario level with caching of invariant matrices and warm starts. Ensemble size is governed by diminishing-returns diagnostics (for example stability of robustness metrics as scenarios are added). Where helpful, a two-stage procedure is used: broad screening followed by a refined ensemble concentrated in discovered vulnerability regions.

Acceptance is included: Registry, bundles, and seeds reproducibly generate the ensemble recorded in the ledger. Compression fidelity holds across scenarios (declared tests pass). Robustness metrics are present in OutlookPack for all runs and correspond to the declared thresholds. For adaptive runs, state continuity checks pass, and all recourse actions satisfy lead-time and eligibility constraints.

This layer shifts emphasis from point forecasts to strategy performance across many plausible futures and from one-off plans to adaptive sequencing. In combination with the feasibility core, it provides the

evidential basis for least-regret recommendations and defensible, monitorable triggers that connect planning to implementation.

6.4.3 Provenance and auditability (R6)

Credible analysis requires a complete and inspectable trace from results back to assumptions. Provenance and auditability are therefore first-class requirements.

Manifests and run ledger: All inputs, configurations, and outputs are accompanied by versioned manifests recording, at minimum: source, method, units, validity window, licence, and a content hash. Scenario bundles are registered as named artefacts with unique identifiers and seeded random states so any ensemble realisation can be reconstructed. A run ledger links each result set to its exact inputs and software context via immutable run IDs, artefact hashes, solver/library versions, and key options (tolerances, gap limits, time limits). Environment capture (e.g., dependency lockfiles or container tags) is recorded to remove ambiguity about execution context.

Acceptance checks (pre-/post-solve): Quality gates are applied before and after optimisation to prevent silent errors and to support reproducible interpretation:

- Schema & referential integrity: field presence, types, units, and key joins (e.g., *site_id, region_code, carrier* labels).
- Temporal alignment: coherence of time indices, representative-day weights summing to annual hours, and wrap-around states for storage.
- Physical consistency: non-negativity, heat/electricity balance closure at the chosen chronology, emissions accounting consistency across site → sub-regional → regional aggregates.
- Post-solve diagnostics: binding-constraint and dual reports, infeasibility certificates where applicable, and reconciliation of ledgers (energy, emissions, cost).

Change management: Schemas and controlled vocabularies are versioned. When a schema changes, migration notes specify how to lift archived artefacts to the new version; deprecations are signalled in advance. Reproducibility is strengthened by deterministic seeds for ensemble generation and by pinning solver settings that materially affect numerical behaviour.

Confidentiality and publication: Where confidentiality applies, redaction rules are recorded in manifests so public artefacts (e.g., aggregates) remain consistent with private sources. Public releases

include the minimal reproducibility bundle: scenario registry, manifests, run ledger excerpt, artefact hashes, and scripts/notebooks that regenerate figures and tables from the recorded outputs.

Acceptance: An independent rerun, using the manifests, ledger, and environment capture, shall reproduce published tables and figures bit-for-bit or within declared numerical tolerances; any failed gate is logged with an explanatory code and remediation note.

This requirement specifies what must be recorded, checked, and linked, rather than how it is implemented, so that independent replication, structured review, and safe iteration are possible as evidence and assumptions evolve.

6.4.4 Evaluation readiness and comparability (R7)

Decision support requires results that are comparable across strategies, epochs, and uncertainty realisations. The evaluation layer therefore standardises metric families, baselines, and reporting conventions so that results can be interpreted consistently across sites, epochs, and ensembles.

Metric families (definitions and scope)

(i) **System cost:** Total and annualised cost (CAPEX annuity, fixed/variable OPEX, fuels and electricity, policy/ETS charges), with the levelised cost of heat (LCOH) reported (where meaningful). Values are in constant currency with the price base and discount rate stated; both totals and deltas to a declared baseline are shown.

(ii) **Emissions:** Annual and cumulative tCO_2e, with optional intensity metrics (e.g., tCO_2e per MWh-heat delivered). Cumulative values are carried across epochs consistent with the temporal scaffold.

(iii) **Adequacy and utilisation:** Unmet heat (MWh and hours), minimum surplus available heat, unit capacity factors, storage cycling statistics, peak demands, and indicative GXP headroom.

(iv) **Transition dynamics:** Installed stock by technology, retirements and additions per epoch, and assets at risk of stranding (capacity with persistently low utilisation).

(v) **Signals and scarcity:** Summary of binding-constraint duals/scarcity rents (e.g., marginal electricity price at GXP, biomass scarcity rent, binding emissions shadow value) over the chosen chronology to explain limiting conditions.

(vi) **Robustness:** Regret, reliability, and satisficing.

Normalisation and aggregation rules

All metrics shall carry units, price base, and (where applicable) discount rate. When temporal compression is used, representative-time weights and reconstruction errors are reported. Additive quantities (cost, energy, emissions) use weighted summation; non-additive quantities (prices, factors, intensities) report the aggregation method (for example energy-weighted mean). Where relevant, per-output normalisations (for example per tonne of product) accompany system totals. Aggregation from site to region is additive for flows and stocks; signals are summarised with appropriate statistics (mean, percentile, and "at-peak" values).

Reporting views

(a) Baseline-versus-strategy deltas by epoch;

(b) Cost–emissions frontiers for shortlisted strategies;

(c) Ensemble summaries (e.g., box/violin plots for cost and emissions; regret–reliability scatter for robustness trade-offs); and

(d) Concise tables of binding constraints with associated duals to explain scarcity and failures.

Internal consistency checks prior to release

Balances close at the chosen chronology; cumulative emissions equal the sum of annual values; levelised costs reconcile with annualised cash flows; hours flagged infeasible are consistent with unmet-heat calculations; and any reported duals correspond to constraints that bind within tolerance. Where compression is used, reconstruction diagnostics (annual/seasonal energy error, peak error, ramp statistics) are included alongside results.

OutlookPack shall include a stable baseline KPI set for every run (cost, emissions, adequacy/utilisation, transition dynamics, signals and scarcity, robustness), tagged by *strategy_id*, *scenario_id*, and epoch. Units, baselines, and aggregation rules are explicit; cross-run comparability checks pass; robustness metrics correspond to thresholds declared in the scenario registry.

The metric panel is designed for external multi-criteria analysis (weights or satisficing sets supplied by users). No implicit weighting is embedded in the core. A stable, transparent bundle of KPIs and robustness measures is produced so that different decision-makers can apply their own preferences while preserving comparability across sites, epochs, and ensembles.

6.5 Decision–Input–Output–Constraint (DIOC) Map

This section fixes, in one place, what each modelled decision is, which inputs it legitimately requires, which outputs it must publish, and which constraints it is allowed to bind and enforce. The map makes decision-visibility explicit, prevents hidden couplings, and provides a clear contract that the future mathematical formulation can honour. It rests on the earlier discussion of spatial tiers and operational chronology and anticipates the later packaging of inputs/outputs as thin-waist artefacts (DemandPack, SignalsPack, OutlookPack).

To keep the contracts tight and defensible, five principles apply throughout:

- **Ownership:** Each constraint is owned at exactly one tier (site, GXP/sub-regional, regional); no duplication.
- **Sufficiency:** A decision may consume only its declared inputs and must publish its declared outputs.
- **Auditability:** Every execution writes a ledger entry (run identifier, artefact hashes, solver options, tolerances).
- **Minimal coupling:** Internal heuristics and solver states remain private; only artefact fields traverse interfaces.
- **Comparability:** Outputs include the core KPI panel (cost, emissions, adequacy/utilisation, transition dynamics, indicative scarcity, and when ensembles run, robustness metrics).

Decision catalogue (compact contracts):

Site investment choice (per site, per epoch). Inputs comprise the site's demand profiles and configuration hints from DemandPack, together with regional policy bands, eligibility flags, indicative electricity prices, emissions factors, and resource envelopes from SignalsPack, plus the prior-epoch state of installed stock and remaining life. Outputs are capacity additions and retirements by technology and epoch, updated stocks, a feasibility flag, KPIs, and a concise binding-constraint summary in OutlookPack; where portfolio changes materially alter profiles, an updated DemandPack may be issued. Constraint ownership is local: technology availability, minimum/maximum output, ramp and temperature feasibility, storage limits, and on-site emissions bookkeeping. Interaction with shared limits occurs only via the light price–quantity exchange when envelopes bind.

Site dispatch and storage operation (hourly feasibility under given signals). Inputs are time-indexed demands (and weights/mapping) from DemandPack, together with time-indexed prices or scarcity

indicators and emissions factors from SignalsPack; storage initial conditions are included where relevant. Outputs are unit dispatch, storage cycles, fuel use, on-site emissions, a feasibility certificate, and dual/constraint diagnostics in OutlookPack; the site also discloses its time-indexed net draws upward as part of its DemandPack. Constraint ownership remains at the site for heat and electricity balances, unit capacity and ramps, storage continuity, and maintenance windows. Only net demands cross the waist; plant-internal detail does not.

GXP/sub-regional access mediation (where employed). Inputs are aggregated site disclosures of net electricity and fuel draws to the relevant GXP or logistics catchment, plus local connection rules and feeder limits from SignalsPack. Outputs are indicative headroom signals, connection timelines, and any curtailment flags returned to affected sites via SignalsPack, with headroom-utilisation diagnostics in OutlookPack. Constraint ownership is limited to envelope checks on connection headroom and simple logistics throughput. No optimisation is performed here; the role is to prevent infeasible local congestion from reaching the regional pass.

Regional electricity envelope (adequacy against aggregated site/GXP draws). Inputs are regional time-series of electricity demand aggregated from DemandPack, along with capacity limits and adequacy rules in SignalsPack, and any emissions interactions where electricity emissions are accounted regionally. Outputs are updated indicative prices or scarcity indicators and time-indexed headroom in SignalsPack, with adequacy/utilisation summaries and binding-constraint duals in OutlookPack. Constraint ownership is regional: adequacy and capacity/headroom limits attributable to the region. When limits bind, signals are returned downward, and iteration proceeds to tolerance.

Regional biomass envelope (deliverability and sustainability across sites). Inputs are seasonal biomass draws by tier aggregated from DemandPack, plus cost–quantity steps and sustainability tags in SignalsPack. Outputs are scarcity rents or indicative prices and available quantities by tier/season in SignalsPack, together with utilisation of tiers (and any admissible unserved biomass) in OutlookPack. Constraint ownership is regional: sustainable availability, seasonal logistics, and environmental conditions. Iteration proceeds with sites when envelopes tighten.

Emissions accounting and compliance (regional and national lenses). Inputs are on-site emissions ledgers from OutlookPack and cap rules, ETS bands, and eligibility information from SignalsPack. Outputs are reconciled regional and cumulative emissions, cap usage, and the implied shadow value when caps bind, recorded in OutlookPack. Constraint ownership sits with the region for budgets and caps; the national wrapper supplies accounting conventions. Signals may reflect cap tightness, but no plant-level constraint is altered here.

Evaluation and reporting (comparability across strategies and scenarios). Inputs are the OutlookPack KPIs, feasibility flags, and concise summaries of binding constraints and duals, alongside manifests and ledger entries. Outputs are the standard reporting views: baseline-versus-strategy deltas by epoch, cost–emissions frontiers, adequacy and utilisation summaries, transition dynamics, and compact tables linking binding constraints to indicative duals; when ensembles are run, regret, reliability, and satisficing are added. This layer owns presentation rules only and introduces no operational constraints.

Exploratory ensemble and pathways orchestration (uncertainty and adaptation wrapper). Inputs are the scenario registry (uncertain parameters with ranges and provenance), exogenous draws encoded in SignalsPack, and the epoch-to-epoch state used for pathway replay. Outputs are scenario labels, regret, reliability, satisficing flags, trigger hit-rates, and failure certificates where applicable, all written to OutlookPack. The wrapper creates no new constraints; it perturbs inputs and sequences runs, leaving operational constraints where they are owned. Discovery is diagnostic and does not modify feasibility logic.

To ensure conformance with these contracts, a run is admissible only if each invoked decision consumes the inputs stated above, publishes the stated outputs, binds constraints solely at its owned tier, and records a complete ledger entry. Any module that requests undeclared fields or attempts to bind a non-owned constraint fails fast at orchestration with a clear error. Where decisions interact, most notably site choices against regional envelopes, coordination proceeds through a light price–quantity exchange with explicit stopping criteria based on balance tolerances, stability of indicative duals, and stability of the binding set.

6.6 Data preparation and packaging

This section translates the decision-centred requirements and chosen scales into traceable inputs and compact data packages that are ready for analysis. Two complementary passes are used. From the top down, a concise register links decisions, constraints, and state variables to the spatial and temporal indices used across the framework. From the bottom up, raw sources are acquired and normalised (units, identifiers, chronology, and provenance are aligned) so that every quantity entering the analysis is auditable.

Operational time is expressed on a defensible chronology (either hourly or a weighted representative set). Compression mappings and boundary conditions are defined once and applied consistently, with diagnostics reported for annual energy, peaks, ramps, and storage cycling. Quality-assurance gates (range and balance checks, completeness, and referential integrity) are applied at each hand-off.

Inputs are assembled into three thin-waist artefacts that decouple data from solvers: DemandPack (site/process characteristics and time-series demand), SignalsPack (contextual inputs such as prices, headroom, and resource envelopes), and OutlookPack (outcomes, diagnostics, and when used robustness labels). Schemas are versioned with required and optional fields distinguished; each artefact instance carries a manifest and a content hash for reproducibility. Profiles allow optional facets to be enabled only when relevant to a decision objective, without changing the core schema.

The end-to-end pipeline "raw → clean → conformant tables → artefacts → runs → diagnostics" is governed by a run ledger and pre/post-acceptance tests. Portability across sites and regions is supported by invariant schemas with local parameterisation, and confidentiality is managed by separating sensitive site fields from publishable aggregates. The proof-of-concept in the next chapter uses the same identifiers, chronology mapping, and artefact schemas at reduced scope.

6.6.1 Decision-centred schema (top-down)

A coherent identifier system and basic governance rules enable reproducible analysis and clear hand-offs between sources, cleaned tables, and artefacts. The objective is to maintain a direct line of sight from decision questions to the data needed for evaluation, avoiding ad-hoc additions while preserving extensibility. A minimum set of indices, manifests, and quality gates anchors every record used later in optimisation and exploratory runs.

The schema organises variables into three classes: decision variables (controllable actions such as technology adoption, sizing, and dispatch), state variables (system memory across epochs such as installed capacity, remaining life, cumulative emissions), and context parameters (exogenous conditions and limits such as prices, headroom, and envelopes). Coherence across modules is obtained by a small, stable catalogue of identifiers.

Each entry links these classes to (i) spatial indices (site, GXP/sub-region, region, national), (ii) temporal indices (operational chronology and decision epochs), and (iii) auditable data fields. To ensure coherence across modules and scales, the register is keyed by a small, stable catalogue of identifiers as shown in Table 6.2.

Table 6-2 Canonical indices and identifiers

Category	Name	Purpose (scope)	Notes
Spatial	*site_id*	Unique industrial facility	Stable across epochs
	gxp_id	Grid Exit Point context	Optional for non-electric carriers
	region_code	Regional aggregation tier	Consistent with NZ usage
Temporal	*epoch_id*	Planning epoch (year)	Aligns with policy steps and state updates
	ts_id	Hourly timestamp	Use when full chronology is used
	rep_id, weight	Snapshot/representative index and weight	Weights sum to annual hours
Technology/ process	*process_id*	Intra-site process	Optional but helpful for audits
	unit_id	Physical utility unit	Maps to *tech_code*
	tech_code	Technology class	Controlled vocabulary
Carrier/fuel	*carrier*	Energy carrier	e.g., heat, electricity
	fuel_code	Fuel type	e.g., coal, residues, pellets
Scenarios	*scenario_id*	Scenario bundle key	Links to uncertainty register
	draw_id, seed	Ensemble member & random number generator (RNG) seed	For reproducibility
Chronology mapping	*cluster_id*	Optional time-cluster label	Optional; only for representative-time clustering
Governance	*version, record_hash*	Change tracking and integrity	With source, method, validity window

For each decision domain (technology, fuel, infrastructure, emissions, policy and governance, temporal evolution), the register records: the decision question; the decision variables indexed by the relevant spatial and temporal keys; the binding constraints at their proper tier (site balances and temperature feasibility at the operational chronology; GXP headroom and biomass envelopes at sub-regional/regional levels; cumulative or annual emissions where relevant); the state updates carried between epochs; the context parameters by tier; and the expected sources (for example RHDD and site documentation for plant and process data; MBIE, EECA, NZ ETS for policy and emissions; Transpower for connection data; NIWA or reanalysis for weather shaping).

Two principles keep the schema compact. First, fields are included only where they evaluate a decision or enforce a constraint at its owning tier. Second, indices are introduced only when aggregation or signalling would otherwise be ambiguous. When chronology is compressed, a single, versioned mapping table links each representative index to the constituent hours so that energy totals, peaks, ramps, and storage states reconcile within stated tolerances.

The register is the logical precursor to packaging. Each row points to the eventual receiving artefact: operations (time-series and configuration) for DemandPack, contextual limits and signals for SignalsPack, and outcomes and diagnostics for OutlookPack. In this way the same decision-centred

schema serves simultaneously as documentation of modelling intent and as the scaffold for portable, implementation-agnostic data packages.

6.6.2 Source acquisition and normalisation (bottom-up)

The process for obtaining, verifying, and transforming raw datasets into conformant data tables aligned to the identifier catalogue and register is fully transparent. The objective is to produce inputs that are traceable, internally consistent, and ready for packaging without committing to any software stack.

Core inputs are drawn from public and site-provided datasets, including plant and process documentation, regional energy transition studies, national energy statistics, grid access and hosting capacity reports, and meteorological series used for demand shaping. For each dataset, a manifest records retrieval date, licence/permissions, native units, temporal coverage, and known caveats. A frozen snapshot (version tag and checksum) is retained so that future re-runs reproduce the same numbers. Heterogeneous inputs are converted to common physical and economic bases before modelling, using a single ruleset applied uniformly across sources (Table 6.3).

Table 6-3 Normalisation conventions

Category	Convention	Notes
Units	SI throughout; power in MW, energy in MWh; fuel mass in t; temperature in °C	One canonical column per quantity
Calorific basis	Lower heating value (LHV) for all fuel energy conversions	Basis recorded per row for transparency
Emissions	CO_2-e factors harmonised to one GWP metric and horizon (e.g., AR5-100)	Factors stored with version and method
Currency	NZD in a fixed base year (e.g., 2020); real terms	Deflators and sources logged in manifest
Electricity prices	Time-series aligned to the operational chronology	Loss factors and charges applied consistently
Time zone	Timestamps stored as local civil time with explicit handling of DST	Mapping to modelling chronology kept in a lookup
IDs and vocabularies	Keys must match Indices and identifiers; carriers/fuels from controlled lists	Rows failing key integrity are rejected

Inputs are reconciled to the operational chronology (hourly or representative-day). Hourly sources are re-indexed to {*ts_id*}; sub-hourly series are aggregated with energy-preserving rules; lower-frequency series are disaggregated using documented shaping factors (e.g., temperature or production calendars). When a representative-day or snapshot set is used, a single mapping table links {*rep_id*} to

constituent hours, with weights that sum to the annual hours. Storage boundary conditions and wrap-around states are specified to prevent segment-edge artefacts.

Every record is tagged with the spatial keys required for aggregation and signalling. Sites are mapped to *gxp_id* using documented feeder/connection data; *region_code* is assigned using a published regional definition. Where multiple plausible mappings exist, the chosen mapping and rationale are recorded, and a sensitivity flag is added for later exploration.

Corrections follow a small set of deterministic rules: (i) range and physics checks (for example, efficiencies in [0,1], non-negative flows); (ii) obvious unit errors (for example, kW vs MW) corrected using diagnostic ratios and logged; (iii) missing values imputed by a ranked protocol (site-specific history → similar-process proxy → engineering estimate), with *impute_flag* and *impute_method* recorded; and (iv) conflicting duplicates resolved by source precedence (evidence strength), with discarded records retained in an audit table.

Records progress from raw to clean to conformant only if they pass the checks in Table 6.4. Failing rows are quarantined with reason codes; runs do not proceed while required gates remain open.

Table 6-4 Quality Assurance (QA) gates for conformant inputs

Gate	Check	Applies to	Outcome on fail
Q1 Keys	Foreign-key integrity for *site_id, gxp_id, region_code*	All tables	Reject row; raise integrity error
Q2 Units	Canonical units	Numeric fields	Convert or reject; log action
Q3 Time	Chronology alignment; sum of weights equals 8,760	Time-indexed tables	Recompute mapping and weights
Q4 Balances	Site annual heat/electricity balances within tolerance	Site aggregates	Flag for reconciliation
Q5 Ranges	Physics bounds and plausibility limits	Technology and process records	Reject or clip per documented rule
Q6 Provenance	*source, method, version*, *valid_from/to*, *checksum*	All tables	Block promotion to “conformant”
Q7 Price base	Currency base-year consistency	Cost and price series	Deflate or inflate; log transform

Each data table carries a compact header (key–value pairs) with source, retrieval date, version, licence, transform steps, and a table-level checksum. Row-level fields include *record_hash* to detect silent changes. A run ledger records which table versions were used, the random seeds applied in any sampling, and the acceptance status of all QA gates.

The bottom-up process yields clean, ID-keyed, unit-harmonised tables for site operations, infrastructure envelopes, emissions and accounting parameters, and demand-shaping series; lookup

tables for temporal and spatial mappings; and a manifest plus ledger entries that enable independent replication. These deliverables are the immediate substrate for packaging into the stable data artefacts used by the analytical layers.

6.6.3 Chronology and compression mapping

This subsection fixes the operative time base for operations and sets a reproducible compression method that maps hourly series to a small, weighted set of representative days (or snapshots) while preserving decision-relevant features. The mapping is declared once per study, versioned, and applied consistently to all time-indexed inputs and outputs. The objective is to reproduce annual energy, seasonal patterns, peak conditions, and ramping behaviour closely enough for technology sizing, adequacy checks, and storage operation without the computational burden of full-year chronology.

Selection of the representative set

The year is stratified by season (for example, Summer, Shoulder, Winter) and day type (Weekday, Weekend). Optional tags (production calendar, maintenance periods) are added only when they materially affect profiles. Within each stratum, daily profiles are reduced to a compact descriptor (daily energy, peak hour, intraday ramp statistics, temperature proxy). A small set of representatives per stratum is then chosen using a simple distance-based routine (for example, k-medoids or quantile picks), with two safeguards:

(i) stress inclusion: the global peak condition is included explicitly either via the actual peak day or a single-hour “stress slice”; and

(ii) diversity control: near-duplicates are avoided by enforcing a minimum separation in descriptor space.

Each representative carries a weight equal to the number of original hours it stands for; weights sum to the annual hour count (8,760 in non-leap years; 8,784 in leap years).

Mapping table (single source of truth)

A single, versioned table governs all mappings from original hours to representatives. At minimum it contains: the original time index (or timestamp), *rep_id*, *weight*, stratification tags (season, day type), *order_in_day* for 24-hour representatives, a *stress_flag*, and a *mapping_version*. Optional fields include a clustering label and tariff band tag to support price audits. Any change to this table invalidates dependent artefacts and triggers regeneration.

Boundary conditions and dynamic fidelity

Representative constructions can distort stateful behaviour if left unconstrained. Three rules avoid this:

(i) Storage closure: for 24-hour representatives, state of charge is cyclic within each representative; a year-level neutrality condition enforces zero net storage energy over the weighted year (with a small tolerance for numerical robustness).

(ii) Ramps: intraday ramp limits are enforced within representatives; ex-post checks compare the empirical distribution of hourly ramps from the original and reconstructed series.

(iii) Peak adequacy: when the representative set does not capture the system peak, a single-hour stress slice (weight = 1) is appended for capacity sizing.

Electricity prices, including time-of-use bands, are aligned by mapping on (*gxp_id*, *rep_id*, *order_in_day*); an optional tariff-period field supports coverage diagnostics. If single-hour snapshots are used, only memory-less technologies are dispatched directly; assets with inter-hour dynamics either require linking constraints over an imposed sequence or a separate adequacy check.

Diagnostics and acceptance targets

Compression is accepted only if explicit tolerances are met: annual energy error ≤ 1%; seasonal energy error ≤ 2% per season; no under-capture of the annual peak (the peak day or a stress slice must be present); ramp distribution similarity within a standard threshold (for example, Kolmogorov–Smirnov distance ≤ 0.10 or Root Mean Square Error (RMSE) ≤ 5% of peak); and storage cycle neutrality within 1% of annual demand. Breaches trigger one or more of: increasing the number of representatives, refining stratification (for example, splitting Winter), or adding a stress slice.

Reporting and reproducibility

Each study includes a short chronology report with representative counts by stratum, mapping hash and version, random seeds (if clustering is stochastic), the sum of weights, and the diagnostics above. The report and the mapping table are referenced in the run ledger so that operational results are traceable to the underlying chronology. For the proof-of-concept, a compact set is used and expanded only if diagnostics fail: three seasons × two-day types × two intensity bands (≈12 representative days) plus a peak stress slice, about 289 reconstructed hours. The same mapping is reused across scenarios and epochs to avoid confounding uncertainty with chronology; if re-mapped, a new mapping version is recorded, and dependent artefacts are regenerated.

6.6.4 Artefact design and external linkages

Industrial process heat decisions require integrating plant-level physics with system-wide constraints but entangling raw data with solver code risks brittleness and opacity. Three standardised artefacts (DemandPack, SignalsPack, and OutlookPack) facilitate upward and downward information flows, ensuring traceability and portability across modules. To standardise communication and keep data independent of solvers, the framework exchanges these small, versioned artefacts. Each has a clear role in the data flow, carries only the fields necessary for feasibility and evaluation, and leaves room for optional detail when finer physics or broader scope are required. This enables reproducibility, portability across sites and regions, and disciplined growth in model fidelity without re-engineering the pipeline.

The consequences of this approach are threefold: (i) every number is traceable to a manifest and version; (ii) the same packages can be re-used with different solvers or in future studies; and (iii) new detail can be added without breaking existing consumers. At a high level, DemandPack carries site/process characteristics and operational time-series, SignalsPack conveys context from shared constraints (prices, headroom, envelopes, allowances), and OutlookPack records outcomes and diagnostics suitable for strategy comparison and robustness analysis. The artefacts are carrier-agnostic and chronology-agnostic: the same shapes apply under full-hour runs or representative-day mappings, provided indices and weights are respected. Extension hooks support integration with electricity-system models (e.g., PyPSA), activity-based planning models (e.g., TIMES-NZ), and lightweight file-driven frameworks (e.g., BlueSky) without altering the core contracts. Table 6.5 summarises each artefact's purpose, core content, producers/consumers, and QA.

Table 6-5 Artefact overview (purpose, core content, producer/consumer, QA)

Artefact	Purpose	Core content (conceptual)	Producers → Consumers	QA/Versioning
DemandPack	Upward: site/process demand & configuration	IDs, operative chronology, heat & electricity demand, tech/config hints, provenance	Pre-processing / site modules → site solvers; regional aggregators	schema version; mapping version; unit & ID checks
SignalsPack	Downward: system context & shared limits	Time-indexed prices/indicative scarcity, GXP headroom, biomass envelopes (epoch), emissions caps, flags	Regional/electric/biomass modules; policy layer → site solvers/dispatch	consistency with regional ledgers; monotonicity/range checks; schema_version
OutlookPack	Results, diagnostics, and evaluation labels	Feasibility flags, KPIs, constraint/dual summaries, capacity stocks, trigger hits, robustness metrics	Site & regional solvers; ensemble orchestrator → reporting; scenario discovery; MCDA	Parent-artefact hashes; fixed metric defs/units; seed record

The guiding principle is small mandatory, generous optional. Required fields ensure feasibility and accounting are always computable; optional fields accommodate richer dynamics (temperature, ramp metadata, storage diagnostics) and cross-model linkages when appropriate. The design follows the identifier catalogue introduced earlier.

DemandPack provides the upward flow of site-specific data, capturing demand profiles and configuration hints essential for site-level optimisation. For instance, it includes hourly heat_MW profiles derived from site processing cycles, incorporating stochastic variability (±10-15%). Table 6.6 outlines its schema.

Table 6-6 DemandPack schema (required and optional fields)

Field group	Required fields (essentials)	Optional fields (illustrative)	Notes (producer → consumer)
Identity & scope	*site_id*; *process_id* / *unit_id* (optional); *carrier*	*temp_class; product_code*	Sites → site solver; preserves process granularity
Chronology	*ts_id* **or** (*rep_id*, *order_in_day*, *weight*); mapping_version	*season, day_type, production_calendar_tag*	Same mapping is reused across all time-series
Demand series	*heat_MW, elec_MW* (time-indexed)	*cooling_MW; steam_quality flags*	Consistent units (MW, MWh via weights)
Configuration hints	*tech_config_id*; min/max output; ramp hints	*maintenance_windows; availability_factors*	Hints are non-binding unless activated in the solver
Provenance	*source, method, version, valid_from*/to, *record_hash*		Enforced by ingress QA gates

SignalsPack handles the downward flow, providing regional context such as prices and constraints to inform site decisions. For Edendale's coal-to-electric transition, it conveys GXP *headroom_MW* limits, preventing infeasible electrification pathways. Table 6.7 details its schema.

Table 6-7 SignalsPack schema (required and optional fields)

Field group	Required fields (essentials)	Optional fields (illustrative)	Notes (producer → consumer)
Scope keys	*region_code* **or** *gxp_id*; *carrier/fuel_code*	*feeder_id*	Regionals → sites; electricity vs biomass handled symmetrically
Prices & scarcity	*elec_price* (time-indexed) **or** *elec_marginal_indicator*; *biomass_price* / *scarcity_indicator*	*loss_factor; network_charge_component*	"Indicator" allows price-free scarcity signalling if desired
Capacity & envelopes	*elec_headroom_MW (time-indexed); biomass_envelope_t (epoch-indexed)*	*connection_queue_flag; reinforcement_eta*	Headroom provides a hard/soft limit according to study design
Emissions & policy	*emissions_cap_tCO2e* (region/epoch); *ETS_band tag;* eligibility flags	*free_alloc_frac; cap_rule_version*	Aligns site accounting with regional caps
Provenance	*source, method, version, valid_from/to, record_hash*		Must reconcile with regional ledgers prior to release

OutlookPack captures results for evaluation, including KPIs and robustness metrics. It aggregates site outcomes for regional analysis, such as total_cost and emissions_annual from Edendale simulations. Table 6.8 presents its schema.

Table 6-8 OutlookPack schema (required and optional fields)

Field group	Required fields (essentials)	Optional fields (illustrative)	Notes (producer → consumer)
Run identity	*run_id; scenario_id; epoch_id; mapping_version; seed*	*draw_id*; label	Keys link to manifests and ensemble registry
Feasibility & ledgers	*feasibility_flag; violated_constraints list; energy/emissions/cost ledgers*	*unmet_heat_MWh; curtailment_hours*	Ledgers are additive across sites; constraints carry magnitudes
KPIs	*total_cost; annualised_cost; emissions_annual; emissions_cumulative; utilisation factors*	LCOH; headroom utilisation; storage cycles	Units and price base declared in manifest
Signals & duals	representative_duals summary (electricity, biomass, emissions)	price/quantity time-series excerpts	For interpretability; not a market model claim
Robustness labels	regret; reliability; *satisficing_flag*	*trigger_hits; pathway_id*	Populated when the exploratory layer is active

Portability follows from fixed schemas and identifiers with local parameterisation. Confidential fields (e.g., unit-level costs or product codes) can be isolated in DemandPack and withheld from publication without breaking site or regional arithmetic; OutlookPack contains publishable aggregates by construction. SignalsPack can expose envelopes and indicative scarcity instead of proprietary tariff structures where required by confidentiality.

External linkages are realised via adapters at the edge; the artefact contracts remain stable, avoiding dependencies on any external tool. For PyPSA-style adapters (electricity focus): DemandPack to PyPSA "load" profiles (per bus or GXP); SignalsPack to time-series constraints (line and transformer capacities via headroom; carrier prices as cost time-series); OutlookPack to extraction of nodal shadow prices and adequacy diagnostics. Mapping from *gxp_id* to buses is one small lookup. For TIMES-NZ-style adapters (timeslices/activity): DemandPack profiles to TIMES timeslice demands; SignalsPack envelopes to bounds on activities or commodities and emissions constraints; OutlookPack KPIs to activity cost and emission summaries. Where TIMES requires coarser timeslices, the chronology mapping generates the corresponding slice weights directly from the same table used internally. For BlueSky-style configuration: The chronology mapping emits the hour, day, year grouping files and year weights used in file-driven frameworks; SignalsPack fields (prices, caps, envelopes) align with the simple CSV/TOML parameter style; DemandPack can be reduced to site-level totals if only lightweight screening is required. Manifests mirror BlueSky's run-config concept: a single source of truth for IDs, versions, and toggles.

Future detail can be added without schema breakage through reserved optional fields; representative hooks are shown in Table 6.9.

Table 6-9 Reserved expansion hooks (illustrative)

Artefact	Hook	Intended use
DemandPack	*temperature_profile, humidity_profile*	Weather-shaping of heat/cooling loads
DemandPack	*utility_limits (min_up/down, start_cost)*	Finer unit commitment representations
SignalsPack	*congestion_rent_proxy, feeder_loading*	Finer electrical network feedback
SignalsPack	*biomass_quality_class, storage_logistics_tag*	Supply-chain and logistics coupling
OutlookPack	*LCA_intensity, water_use*	Life-cycle and resource metrics
OutlookPack	*VoI_tag*, *data_gap_flag*	Feedback from robustness/VoI studies to data governance

Each artefact carries a compact manifest (key–value header) with *schema_version*, *mapping_version*, source, method, price base and units, and content hashes. The run ledger stores the tuple of artefact

hashes, solver options, and seeds used in a run so that any OutlookPack can be traced exactly to its DemandPack and SignalsPack parents.

What is standardised is not the internal method of a module but the contract at its boundary. Every artefact declares what it represents, how it was produced, which units and assumptions it carries, which acceptance checks it has passed, and what role it plays in the decision comparison; an artefact that fails its checks does not propagate downstream. DCM is therefore model-agnostic within modules and contract-governed at their boundaries. The internal method appropriate to a module depends on the data available, the maturity of the relevant modelling tools, and the decision relevance of the consequence it carries (Table 6.10). A higher-fidelity method can replace a simpler one without altering its neighbours, provided it emits artefacts that satisfy the same contract.

Table 6-10 Table 6.10 OutlookPack schema (required and optional fields)

Data or model maturity	Internal method	Artefact carried	Status in this thesis
Annual site totals only	Synthetic hourly demand profile	DemandPack	Used in the PoC
Unit constraints and operating logs	Hourly LP or block-wise commitment dispatch	Dispatch ledgers; ΔE	Implemented, not reported
Process-level, temperature-resolved data	Process simulation; heat-pump COP curves	DemandPack facets (temp_class, steam_quality)	Future extension
Regional network data	Headroom screen, then power flow (e.g., PyPSA)	SignalsPack; interface outcomes	Screen used in the PoC
Biomass logistics data	Delivered-cost multiplier, then supply-chain model	SignalsPack (biomass fields)	Multiplier used in the PoC

Together, these design choices keep interfaces small and stable, support reproducible studies across diverse contexts, and enable external validation or co-simulation without compromising the tractable, decision-centred stance of the framework. Building on these artefacts, the next subsection details the computational workflow for ensemble simulations.

6.6.5 Data flow and governance

A single, disciplined pipeline carries data from sources to results. Each transition is explicit, versioned, and gated by checks so that outputs are always traceable to their inputs and assumptions. The flow proceeds from raw acquisition to clean and conformant tables, into versioned artefacts (DemandPack, SignalsPack), through the deterministic core and any site–region coordination, and finally to the OutlookPack and reporting layer. Optional ensemble and discovery runs are orchestrated on top of the

same contracts. Caching and invalidation are governed by content hashes and schema versions so that speed gains never compromise reproducibility.

The standard pipeline consists of the following stages:

1. **Acquisition → raw snapshot.** Source files are fetched and frozen with a manifest (source, licence, retrieval date, native units, coverage, checksum). For example, RETA (RHDD) datasets for region are captured here.

2. **Normalisation → clean tables.** Units, currencies, calorific bases, time zones, and identifiers are harmonised using a single ruleset; all transforms are logged in table headers.

3. **Conformance → conformant tables.** Keys are aligned to the identifier catalogue; chronology is mapped (hourly or representative set); records must pass QA gates before promotion.

4. **Packaging → artefacts.** Conformant tables are assembled into DemandPack and SignalsPack, each stamped with *schema_version*, *mapping_version*, and a content hash.

5. **Deterministic core runs.** Site and regional modules consume the packs; solver options and tolerances are recorded.

6. **Aggregation → OutlookPack.** Results, ledgers, constraint (dual) summaries, KPIs, and (if applicable) robustness labels are compiled and tagged with *run_id*, *scenario_id*, and the artefact hashes.

7. **Ensembles & discovery.** The orchestrator repeats steps 5–6 across scenarios (pathways), caching invariant matrices and reusing warm starts to handle uncertainties like biomass volatility (±15%).

8. **Reporting & archival.** Figures and tables are generated from OutlookPack; manifests and ledgers are archived for reproduction.

Every table and artefact includes a compact key–value manifest (*schema_version*, *mapping_version*, *source*, *method*, units and price base, *valid_from/to*, checksum). The run ledger binds inputs to outputs: *run_id*, *scenario_id*, *draw_id*, seed, solver version, library versions, tolerances, and artefact hashes. Re-running by *run_id* reconstructs the environment deterministically (subject to seed).

Invariant matrices and factorisations are cached per (*schema_version*, *mapping_version*, *model options*). Invalidation triggers automatically on changes to chronology, schemas, keys, or hashes, for example updating mappings regenerates time-indexed artefacts. Figure 6.4 illustrates the pipeline, with gates and caching highlighted.

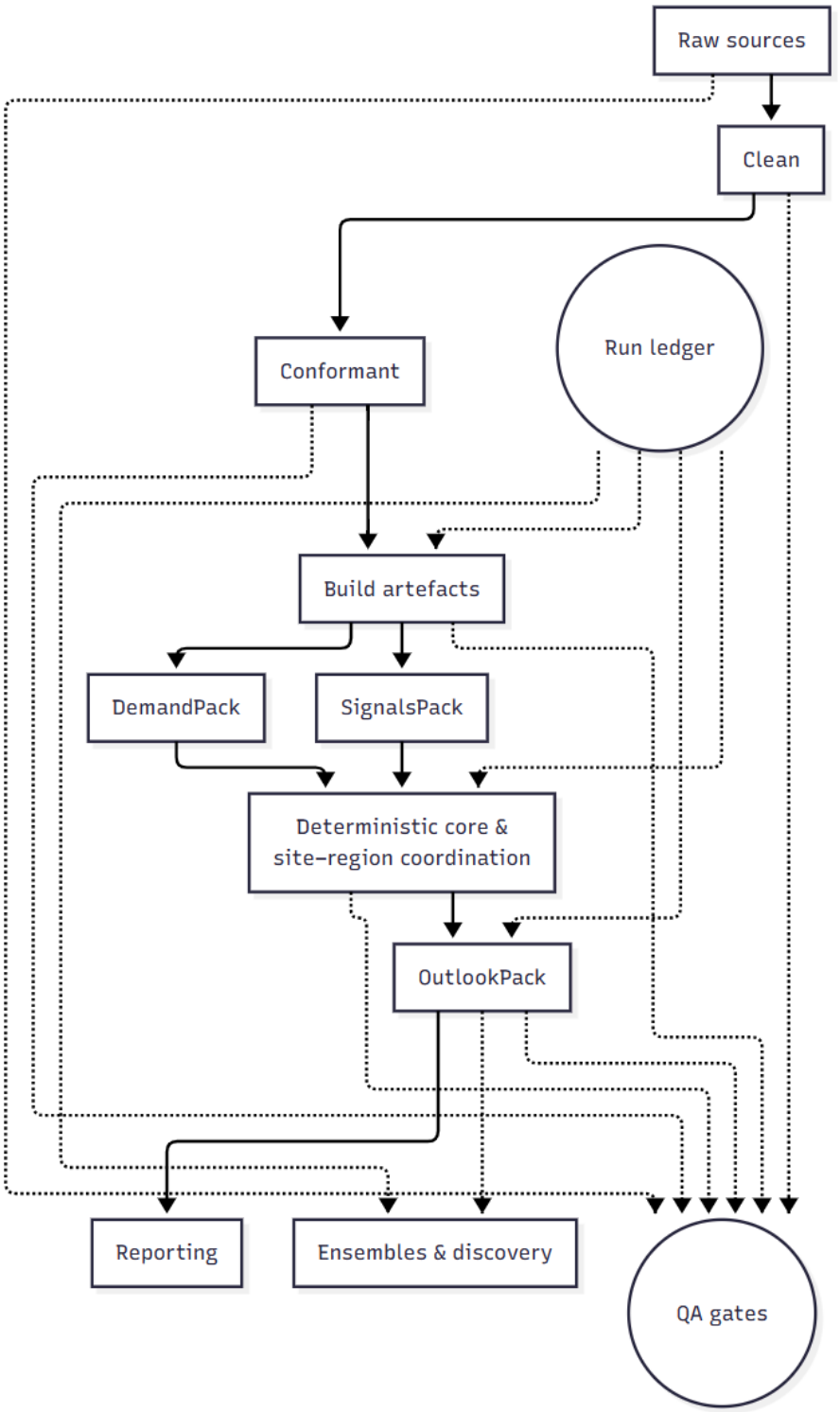


Figure 6.4 Data pipeline in a standard run

Progress is gated by stage-specific tests to maintain data quality. Table 6.11 summarises checks, tolerances, and failure outcomes.

Table 6-11 Acceptance gates

Stage	Check	Tolerance/target	Outcome on fail
Pre-ingest	Schema conformance (fields, types, units)	100% pass	Reject rows; log reason
Pre-ingest	Identifier integrity (*site_id*, *gxp_id*, *region_code*)	No orphan keys	Reject rows; raise integrity error
Pre-ingest	Provenance present (source, version, checksum)	Required fields present	Block promotion to "clean"
Post-normalise	Unit & currency base consistency	Single canonical unit/base	Convert/deflate; log transform
Post-normalise	Chronology alignment (weights sum)	∑ weights = 8,760 (8,784 leap)	Recompute mapping; stop if mismatch
Post-normalise	Physics bounds (non-negativity, 0≤η≤1)	100% pass	Reject/clip per rule; flag table
Pre-run	Artefact presence & hashes recorded	Both packs present; hashes valid	Abort run; regenerate packs
Pre-run	Envelope sanity (headroom ≥0, caps ≥0)	100% pass	Abort run; repair SignalsPack
Post-run	Balance closure (heat & electricity)	≤ 10^{-6} relative (or stated tolerance)	Mark run infeasible; report constraints
Post-run	Emissions ledger reconciliation (site → region)	Exact match within rounding	Flag; recompute aggregation
Post-run	Reproducibility (seed, solver versions pinned)	Recorded in ledger	Block archival until recorded
Post-ensemble	Coverage & stability of robustness metrics	Metrics stable within ± ε; minimum scenario-bin occupancy achieved	Extend ensemble and/or refine sampling; report limits

Schemas and identifiers remain fixed; local parameterisation occurs via artefact content. Sensitive fields can be redacted in DemandPack without breaking balances, while OutlookPack is publishable by design. Schema changes are versioned with migration notes; deprecated fields have a grace window. Chronology reports are archived, preserving links to timebase.

For ensembles, robustness metrics (e.g., regret, reliability) must stabilise within ±5% under resampling, with minimal occupancy across uncertainty bins (e.g., ETS prices, demand variability). Failure prompts automated extension or flagged limitations. In the Edendale case, this ensures reliable stress-testing against grid constraints.

This governance supports reproducible, uncertainty-aware modelling, bridging site granularity with regional planning.

6.7 Operational specification: capabilities, layers, and coordination

Industrial process-heat transitions are made at sites, constrained by regional resources, and scrutinised by national policy; without a clear articulation of capability and scope, models either over-promise with brittle detail or under-inform with coarse averages. The framework therefore requires explicit statements about what level of description is necessary to answer decision questions, how local physics and regional constraints inform one another, and where uncertainty is handled without false precision. These justifications are not administrative niceties: they are the safeguards that keep the analysis decision-relevant, auditable, and extendable.

Site decisions (boiler retirements, electrification timing, storage sizing) are sensitive to temperatures, ramps, and headers, yet their feasibility is also shaped by GXP headroom, biomass logistics, and emissions envelopes. Stating the analytical span prevents category errors, where the site layer carries the thermodynamics and operational constraints, the regional layer carries shared resource limits and scarcity signals. A thin-waist exchange (DemandPack, SignalsPack, OutlookPack) is adopted because it lets each side evolve internally while communicating what the other actually needs to know. Iterative coordination (Gauss–Seidel style) is chosen to mirror real planning, where prices and headroom inform sites, revised demands feed back to the region, and convergence marks a coherent plan.

Decisions unfold over epochs, operations over hours. Without nesting operational resolution (hourly or representative-time with declared weights) within decision epochs, pathway analysis collapses into static snapshots or, conversely, becomes intractable. Declaring this nesting makes path-dependence observable (stock evolution, headroom depletion, emissions ledgers) while keeping runs tractable. Temporal compression is permitted only behind diagnostics (energy, peaks, ramps) so that simplification is earned, not assumed.

Deterministic optimisation is indispensable for feasibility and cost accounting, but it does not, by itself, answer the policy question: “What remains satisfactory across many plausible futures?” An exploratory layer (ensemble sampling, robustness metrics, signposts and triggers) sits alongside the deterministic core to prevent over-interpretation of point solutions and to direct modelling effort where it buys the most decision value. This division of labour clarifies why some inputs are ranges, why results are distributions, and why learning (VoI/VoL) is part of the specification rather than an afterthought.

Capability without discipline breeds irreproducibility. Provenance-rich artefacts, manifests, acceptance tests for compression and coordination, and adapter contracts for external tools need to be specified, so that every figure and table can be traced to its assumptions, units, and code versions.

Interoperability is defined narrowly (through adapters), allowing validation against PyPSA or alignment with TIMES-style outlooks strengthens, rather than distorts, the internal logic.

6.7.1 Analytical capability specification and scope

The environment must be explicit about what it can do, at what resolution, and with which guarantees, independent of any particular solver, dataset, or case. The environment is designed to (i) evaluate industrial process heat pathways under deep uncertainty; (ii) couple site-level engineering feasibility to regional resource and network envelopes; and (iii) report decision-relevant metrics for policy and investment.

These capabilities are organised along three axes.

- **Scale:** Computations must remain coherent from the site to the sub-regional and regional levels, with vertical coordination via standardised data artefacts and iterative signal exchange.
- **Time:** Operational resolution (hourly or representative-time with stored weights) must be nested within decision epochs (e.g., 2020, 2025, 2030, …) to support pathway evaluation and state tracking (stocks, headroom, ledgers).
- **Decision logic:** Deterministic optimisation establishes feasibility, costs, and emissions under stated conditions; an exploratory layer tests robustness, adaptation, and learning across wide uncertainty ranges.

Within scope are multi-carrier interactions linking heat, electricity, and biomass, with explicit thermodynamic and capacity constraints at the site and resource (network) envelopes at regional scale. The environment must support sequential pathway evaluation across epochs, including stock evolution, emissions ledgers, and network headroom states, so that path dependence and timing trade-offs are visible. Assumptions and transformations must be transparent, where every parameter carries units, provenance, and schema versioning. Internally, the environment adopts a graph-based representation suited to tractable optimisation; externally, it communicates through thin-waist interfaces, so internal detail can evolve without breaking inter-module coupling. Interoperability is achieved via adapters that map these artefacts to external ecosystems (e.g., electricity models or policy outlook tools) while preserving methodological independence and avoiding lock-in.

Out of scope (at this stage) are behavioural and agent dynamics and endogenous macroeconomic feedbacks, which are treated as external inputs or future extensions; detailed market-clearing mechanisms and intraday bidding strategies, since the environment consumes prices and scarcity

signals rather than clearing markets; and fine-grained logistics co-optimisation beyond regional biomass and electricity envelopes, which may later be integrated through side-modules or adapters. Declaring these exclusions avoids false precision and preserves tractability while leaving clear hooks for future development.

Quality and tractability are governed by explicit policies. Temporal compression (e.g., representative-time clustering) is permitted where diagnostics demonstrate annual energy reconstruction errors no greater than 1–2%, peak preservation within 5–10%, and acceptable fidelity of ramp statistics; these diagnostics are recorded alongside the artefacts that use them. Modularity and replaceability are mandatory, as sub-modules (such as COP maps or boiler efficiency curves) may be swapped provided interface contracts are honoured. Reproducibility is enforced through run manifests that fix calendars, identifiers, units, random seeds, and configuration hashes, with full data lineage stored alongside results. Execution must scale from workstation-level screening to larger ensembles without modification of external interfaces.

A minimum reporting set ensures decision relevance. Operational outputs include carrier- and unit-level energy use, binding constraints, and (where available) dual/scarcity signals for the next iteration. Economic and environmental outputs include total and levelised costs (ETS-inclusive components), absolute and intensity-based emissions, and declared compliance indicators. Robustness/adaptation outputs include regret and satisficing metrics across ensembles, the incidence of signposts/triggers, and pathway feasibility flags. All results are linked to their underpinning assumptions through the governance layer (the cross-cutting provenance, QA spine comprising per-table and per-artefact manifests, the run ledger, controlled schemas, identifiers and acceptance tests QA gates).

Finally, the specification recognises an evolution path without prescribing an implementation roadmap. Near-term priorities include richer site data, tariff structures, thermal-storage dynamics, and improved biomass envelopes. Medium-term developments may add endogenous retrofit timing (integer flags), semi-Markov project-delay representations, and extended uncertainty governance. Longer-term extensions may incorporate behavioural and multi-agent layers, logistics co-optimisation, and ingestion of monitored (IoT) data for digital-twin updates. All extensions must remain compliant with thin-waist schemas and the acceptance principles above.

6.7.2 Functional layers and iterative coordination

The environment separates concerns across three functional layers and coordinates them through explicit, auditable exchanges, so that engineering realism, system consistency, and decision relevance are maintained without unnecessary entanglement. The deterministic core captures the engineering

and economic logic at the site and regional scales. It enforces energy–mass balances, temperature and pressure feasibility, capacity and ramp limits, and declared policy constraints. Internally, sites are represented on a graph (units, headers, stores, carriers), enabling tractable LP/MILP formulations for dispatch and technology use under stated conditions. Regional modules impose shared envelopes (e.g., GXP headroom, biomass availability, emissions caps) and compute scarcity information (dual variables) associated with binding constraints. The outputs of the core are physically consistent dispatches, cost and emissions ledgers, and the shadow prices that quantify marginal scarcities.

An exploratory layer probes uncertainty and adaptation without altering physics. It builds ensembles of futures via design-of-experiments or Monte Carlo sampling, evaluates candidate strategies across those futures, and computes robustness metrics (e.g., regret, satisficing rates, adequacy). It also identifies signposts and triggers for Dynamic Adaptive Policy Pathways, thereby specifying when a strategy should adapt as conditions evolve across decision epochs. In parallel, a governance and provenance layer guarantees transparency and reproducibility: configuration manifests (identifiers, calendars, units, seeds), schema versions for artefacts, dependency tracking, and acceptance diagnostics are recorded so that every inference can be traced to its assumptions.

Vertical coordination between site and regional scales proceeds by message passing over the thin-waist interfaces. Sites export DemandPack (time-resolved heat and electricity demands and relevant metadata) to the regional level; regional modules return SignalsPack (time-resolved prices, marginal emissions, headroom and scarcity signals) to the sites; OutlookPack aggregates outcomes for decision evaluation and reporting. Coupling follows a Gauss–Seidel logic: the region solves on current site demands, issues updated signals, and only those sites materially affected re-optimise; the process repeats until the system is consistent. To enhance stability where discrete choices appear, updates may be damped by blending new signals with previous iterates, and the coordinator can operate asynchronously so unaffected sites are not re-solved.

Convergence is declared on simple, transparent tests that preserve tractability: (i) signal stability, with changes in key signals (e.g., locational electricity scarcity price, biomass rent, available headroom) below prescribed tolerances, accounting for representative-time weights where compression is used; (ii) constraint closure, with no violations of regional envelopes or site-level feasibility; and (iii) ledger reconciliation, with energy and emissions accounts balanced within declared tolerances across layers and time representations. If iterates stall or cycle, the coordinator applies damping, tightens slacks, fixes low-impact levers temporarily, and records the configuration as unstable, if necessary, before proceeding with the remaining ensemble, thereby protecting throughput.

Horizontal coordination governs the exchange between deterministic evaluation and exploratory assessment. The exploratory layer first parameterises an ensemble by sampling the Uncertainty Register (ETS trajectories, tariff structures, demand growth, biomass envelopes). For each draw, the deterministic core executes the vertical solve to produce feasible dispatch, costs, emissions, duals, and feasibility flags. Results are then aggregated to compute robustness indicators and to perform scenario discovery, revealing combinations of uncertainties that drive success or failure and informing where additional fidelity or data collection would have the highest Value-of-Learning. Updated signposts and trigger rules are recorded for subsequent decision epochs, ensuring that adaptation is explicit and auditable rather than ad hoc.

Execution is engineered for reproducibility and speed. Site and regional solves are warm-started and memorised using keys derived from inputs and active signal sets; factorisations are reused where possible; ensembles are parallelised horizontally; and vertical passes prefer asynchronous updates to avoid needless precomputation. Where representative-time compression is employed, both inputs (DemandPack) and outputs (SignalsPack, OutlookPack) carry weights, and reconstruction diagnostics (annual energy, peak, ramp statistics) are embedded alongside results. Each run writes a manifest with schema versions, solver and package versions, seeds, calendars, units, and hashes of inputs and artefacts, enabling exact re-execution or forensic comparison.

Interoperability with external ecosystems is achieved by adapters that translate thin-waist artefacts to and from external data models (for example, importing national scarcity signals into SignalsPack, or exporting aggregated demands and KPIs from OutlookPack to sector-wide outlook tools). Adapters do not alter internal semantics; they provide context and validation without creating hidden dependencies or lock-in. Overall, the layered design clarifies who decides what, when, and with which information: engineering feasibility resides in the deterministic core; uncertainty exploration and adaptation reside in the exploratory layer; integrity and transparency are enforced by governance. Vertical Gauss–Seidel coupling aligns local feasibility with regional constraints, and horizontal iteration aligns point estimates with robust, adaptive decision insight, delivering a coordinated yet tractable environment for site-to-region industrial process-heat planning.

6.7.3 Complexity management and compression policies

This section defines how complexity is governed so that the environment remains tractable without sacrificing decision relevance. Compression is treated as a controlled transformation (declared, diagnosable, and reversible), rather than an ad-hoc shortcut. Policies apply to time, space, technology detail, and uncertainty, and are enforced through acceptance tests recorded alongside the artefacts

that use them. The guiding principles include (i) preserve the physics and constraints that bind choices; (ii) retain the signals that drive those choices (prices, scarcities, temperature lifts); (iii) record approximation error and provenance; (iv) fail safe by escalating fidelity when tests are not met.

The native operational resolution is hourly. Where justified, representative-time sets may be constructed via seasonal stratification and clustering using features that matter for decisions (e.g., heat load, electricity tariffs, ambient temperature, production calendars). Every representative period carries a weight that is written into DemandPack and propagated to SignalsPack and OutlookPack. Acceptance tests are mandatory and stored with the artefact, where annual energy reconstruction error should not exceed 1–2%; peak preservation should remain within 5–10%; and ramp distributions should lie within agreed bounds (for example, the 95th percentile ramp within ±10%). Extremes (historical peaks, low-renewable windows) are pinned to prevent dilution by clustering. If tests fail for a site or epoch, the policy is to increase cluster count, stratify more tightly (e.g., by season, production phase), or fall back to full-hourly for the affected slice only.

Spatial complexity is managed by hierarchical aggregation rather than global lumping. Sites may be grouped into archetypes only where constraints and choices are genuinely similar; otherwise, they remain distinct optimisation units. At the regional level, GXPs and feeders are aggregated to the minimal set that preserves binding constraints (hosting limits, loss structure, nodal price differentials). Aggregation is prohibited around known bottlenecks; if a GXP headroom constraint binds in trial runs, that node is retained at full fidelity. All mappings from detailed to aggregated sets are versioned and stored with the run manifest.

Technology representations follow a “fit-for-purpose” ladder. COP curves and boiler efficiencies may be linearised with a small number of breakpoints for screening; additional segments are enabled when a unit is shortlisted for investment or when regret analysis indicates sensitivity to its non-linearities. Integer flags (on/off, build/no-build) can be relaxed during broad screening and reinstated for shortlists or validation. Minimum up/down and start-cost logic is out of initial scope, but a hook is declared for later inclusion. Each change in model order is captured in the manifest so that comparisons remain auditable.

Ensemble size is controlled through design-of-experiments and scenario reduction that preserve tails and interactions. Latin Hypercube or Sobol sampling creates an initial superset; outcome-space clustering and scenario discovery then select a reduced set of “distinctive futures” for robustness and reporting. Parameters with consistently negligible influence are fixed (with justification) to avoid scenario explosion. Where Value-of-Learning indicates that a parameter drives regret, its resolution is increased in subsequent batches rather than densifying the entire space indiscriminately.

In the Gauss–Seidel exchange, only sites materially affected by updated signals are re-solved (asynchronous coordination). Numerical oscillations are addressed with signal damping (convex blending of successive signal vectors) and, if needed, temporary fixation of low-impact discrete levers. Convergence checks (signal stability, constraint closure, ledger reconciliation) are applied with the same weights used in temporal compression to avoid bias.

Scarcity signals derived from compressed runs are treated as weighted proxies and are never generalised without verification. Where duals inform investment or pathway triggers, a spot-check policy applies that re-solve targeted epochs at full-hourly resolution to verify that binding status and marginal values persist. Discrepancies above declared tolerances trigger either an increase in representative-period count for the affected season or a local full-hourly override.

Execution proceeds through three standard tiers that can be invoked per module or per site:

(i) **Screening:** Compressed time, relaxed integers, coarse technology segments for rule-out and scoping.

(ii) **Shortlist:** Tighter compression, reinstated integers for shortlisted assets, refined envelopes for portfolio comparison and preliminary robustness.

(iii) **Validation:** Full-hourly for critical windows, full integer treatment on shortlisted decisions, and no aggregation at binding nodes for decision confirmation and reporting.

Pre-solves tighten bounds; units are standardised; big-M parameters are minimised and logged; scaling is applied where condition numbers warrant it; and solver tolerances are fixed by policy. Each run writes a manifest with schema versions, solver and package versions, seeds, calendars, units, and hashes of inputs and artefacts, enabling exact re-execution or forensic comparison. Where representative-time compression is employed, reconstruction diagnostics (annual energy, peak, ramp statistics) are embedded alongside results in OutlookPack.

Compression is disallowed where it would mask the causal mechanism under study, such as near transition thresholds (e.g., headroom signposts), during commissioning or retirement years, or when hourly tariff structures are the primary driver of dispatch. In these circumstances, local full-hourly overrides are mandatory.

Collectively, these policies make complexity an object of design rather than drift. They protect the causal pathways that matter for decisions, keep computational cost proportional to value, and ensure that any simplification is transparent, tested, and (when required) undone.

6.8 Governance, assurance, interoperability, and roadmap

Effective decision support does not end with a sound optimisation core; it relies on governance that makes results trustworthy, interfaces that make them portable, and a development path that keeps the environment durable. This section establishes that backbone. First, it specifies how correctness, credibility, and traceability are assured through disciplined quality controls and full provenance. Second, it formalises how the environment exchanges data and signals with external ecosystems via thin-waist adapters to validate, inform, or consume results without dictating internal design. Third, it introduces a living register of assumptions, risks, and declared limitations to prevent hidden judgements and to guide proportionate mitigation and targeted data collection. Finally, it sets out an extensibility roadmap that sequences enhancements from a stable core to richer coupling and, in time, cross-sector scaling, always under compatibility and acceptance gates. Together these elements turn a modelling architecture into a reliable decision environment: auditable in method, interoperable in practice, and evolvable by design.

6.8.1 Quality assurance, validation, and transparency

The decision environment earns trust by being correct, truthful, and auditable. Quality assurance governs how analyses are produced (process discipline); validation tests whether results respect physics and evidence (fitness for purpose); and transparency ensures that every inference can be traced and reproduced. The principles below apply to every run, irrespective of solver, dataset, or case.

The scope rests on four pillars:

(i) verification of implementations against the declared mathematical specification;

(ii) validation of outputs against conservation laws, envelopes, and reference data;

(iii) reproducibility through complete provenance and deterministic pipelines; and

(iv) openness of assumptions and approximations, including known limitations.

Tests are automated where feasible and recorded with each run’s artefacts.

All inputs pass schema validation (required fields, types, controlled vocabularies), range and consistency checks (non-negativity, bounds, monotonicity), and referential integrity (IDs/foreign keys). Any imputation or cleaning records method, parameters, and affected records in the artefact metadata.

Source, manifests, and schemas are version-controlled; changes occur only via reviewed pull requests. Automated tests cover utility functions (e.g., profile shaping, COP interpolation), I/O contracts

(schema, units, value ranges), and a minimal end-to-end integration run. Runtime environments are pinned (lockfiles; container image digest). The declared mathematical programme is checked via unit tests on toy systems with known optima.

Solves report feasibility residuals (mass and energy balances, header feasibility), constraint violations, and standard optimality diagnostics. Dual/shadow prices are sanity-checked for sign, magnitude, and complementarity consistency. For MILP, integrality gaps are set by policy per fidelity tier (e.g., ≤0.5% for validation). IIS (irreducible infeasible set) diagnostics are archived for infeasible instances.

Results must satisfy conservation, nameplate capacities, temperature/pressure feasibility, ramp/availability rules, and regional envelopes (e.g., GXP headroom, biomass supply). Where reference series exist (annual totals, seasonal shapes, emission factors), KPIs are compared within declared tolerances; discrepancies prompt recalibration or explicit justification. When representative-time compression is used, reconstruction diagnostics (annual energy, peak, ramp statistics) accompany the artefacts.

The Uncertainty Register records ranges, sources, and (where appropriate) distributions. Design-of-Experiments settings (Latin Hypercube or Sobol), seeds, and coverage metrics are logged. Robustness outputs include satisficing rates against declared thresholds, regret distributions, and scenario-discovery summaries (distinctive futures). Where Value-of-Learning indicates high sensitivity, subsequent batches increase resolution selectively on influential parameters rather than densifying the entire space.

In Gauss–Seidel exchanges, convergence requires: (i) signal stability (changes in key signals, weighted by time-slice weights, below thresholds); (ii) constraint closure (no regional-envelope violations); and (iii) ledger reconciliation (energy/emissions balance across site ↔ region). To suppress oscillation, signal damping and asynchronous updates are applied; stalled cases are quarantined and escalated to higher fidelity.

Every run writes a manifest capturing: schema versions; solver and package versions; calendars and time-zone handling; units; seeds; compression settings; and hashes of inputs and outputs (DemandPack, SignalsPack, OutlookPack). Artefacts include their own metadata blocks (assumption registry, source citations, quality scores, valid_from/to). Three reproducibility tiers are supported: exact (bit-for-bit using the recorded container), analytic (numerically equivalent within tolerances on a different machine), and conceptual (methodologically equivalent with alternative but documented datasets). For policy-facing outputs, domain experts review seasonal shapes, tariff responses, and

capacity utilisation against operational knowledge. Review outcomes are logged alongside the manifest.

Assumptions are explicit, machine-readable, and collated in an Assumption Register linked to each result. All approximations (e.g., linearisation segments, relaxed integers) are disclosed alongside their acceptance tests. Significant figures reflect data and model uncertainty; spurious precision is avoided. Each figure or table references the manifest ID and artefact hashes from which it was produced. A *change log* records model and assumption revisions with semantic versions and notes on backward compatibility of artefacts.

The environment supports “replication packs”: frozen manifests plus minimal data slices allowing third parties to re-run selected experiments. Where suitable, comparator protocols are declared (e.g., alignment with established national models for selected metrics); differences are explained by scope and boundary conditions, not hand-tuned parameters. Synthetic benchmarks (toy systems with known optima) are maintained to detect regressions in solvers or formulations.

Failures of verification, validation, or transparency tests are first-class outcomes. Results are flagged, quarantined, and excluded from synthesis dashboards; an issue record is opened with context (inputs, seeds, solver logs), and the escalation path is invoked: increase fidelity, relax compression, refine data, or revise formulation. Only when acceptance checks pass are runs admitted to OutlookPack summaries and decision reports.

Sensitive or proprietary inputs are segregated, minimised, and where possible replaced by ranges. Access controls apply to raw data; only derived, non-identifiable aggregates enter public artefacts. Governance roles (maintainer, reviewer, data steward) and their responsibilities are defined; sign-off is required before releasing policy-facing outputs. FAIR data principles guide publication of artefacts and documentation.

Together, these measures ensure that results are not only numerically correct but also explainable, reproducible, and fit for decision-making.

6.8.2 Interoperability and adapters

This section specifies how the environment exchanges data and signals with external modelling ecosystems, principally PyPSA (network-optimisation, graph-based electricity systems) and TIMES-NZ (long-horizon, technology-rich energy system planning), without diluting the framework’s decision orientation or thin-waist discipline. Interoperability is achieved via adapters, which are narrow, versioned translators that map our canonical artefacts (DemandPack, SignalsPack, OutlookPack) to and

from the external model's native schemas, units, and temporal structures. Adapters isolate heterogeneity at the edges, so the internal architecture, data governance, and verification policies remain stable.

Interoperability is **(i) contract-based:** adapters are defined by machine-readable interface manifests and unit conventions; **(ii) non-intrusive:** no external tool dictates internal structure; **(iii) bidirectional:** imports and exports both carry provenance; **(iv) loss-aware:** aggregation, compression, or discretisation steps are explicit and reversible where feasible; and **(v) auditable:** every exchange writes a manifest with version hashes, mapping rules, and acceptance diagnostics. Adapters are optional; they enable three modes of use: validation (cross-check), soft-linking (iterative exchange of signals), and alignment (harmonised reporting). Each adapter implements two pure functions:

- Writer (canonical → external): *write_<tool> (DemandPack, SignalsPack, manifest) → external_input*.
- Reader (external → canonical): *read_<tool> (external_output, manifest) → SignalsPack | OutlookPack*.

Manifests define field mappings, unit conversions (e.g., MW to t/h steam via enthalpy at header conditions; MWh to GJ), temporal transforms (hourly to representative periods to TIMES timeslices); spatial transforms (site/GXP codes to external region/bus IDs); sign conventions, price and emission semantics, and acceptance tests (round-trip checks, reconstruction errors). Manifests live with the run ledger and are versioned alongside artefact schemas.

PyPSA interoperability leverage PyPSA's graph-based dispatch and nodal pricing to (a) validate electricity-system interactions implied by site transitions and (b) generate locational signals (LMPs, curtailment, headroom) that populate SignalsPack for Gauss–Seidel exchanges.

Canonical → PyPSA (Writer)

- Map each site to one or more buses (electricity, and—if used—heat buses) using region/GXP codes.
- Export DemandPack electricity time series as Load at the connected bus; export flexible electrified heat as Link sinks (for HTHP/E-boiler, electricity → heat) with efficiency or COP curves linearised into PyPSA piecewise segments.
- Represent storage (if any) as StorageUnit/Store; represent exogenous headroom as Line/Transformer limits or GlobalConstraint bounds.

- Apply temporal compression via snapshot weights when using representative periods; write weights to a PyPSA *snapshot_weightings* structure (declared in the adapter manifest).
- For emissions either embed carrier-level emission intensities and GlobalConstraint caps or pass as cost adders consistent with our ETS parameterisation.

PyPSA → Canonical (Reader).

- Import LMPs, congestion components, curtailment, and utilisation at relevant buses; aggregate (or disaggregate) to our time basis using declared weights; populate SignalsPack fields: *elec_price*, *headroom_MW*, *elec_CO_2*, optional *scarcity_flag*.
- For validation mode, import PyPSA dispatch of links (HTHP/E-boiler) and compare to site-module dispatch under identical tariffs; record discrepancies and acceptance diagnostics in OutlookPack.

In iterative runs, the site module uses updated SignalsPack (prices and headroom) to re-optimise flexible heat supply; aggregated electric demand is re-written to PyPSA; iterations damp updates until signal change and constraint closure meet convergence thresholds.

PyPSA's electricity focus means process heat is represented via links and loads; temperature and pressure feasibility remains enforced internally by the site module. Snapshot clustering induces dual-value bias; our duals-under-compression policy mandates spot checks at full-hourly resolution where signals inform investment triggers.

TIMES-NZ interoperability by comparing and aligning long-horizon technology portfolios and system costs with an established national planning model; optionally exchange price/emission trajectories or demand storylines for scenario alignment.

Canonical → TIMES-NZ (Writer)

- Aggregate DemandPack to TIMES timeslices (season/day/night or model-specific structure) with explicit weights; map site/GXP region codes to TIMES regions.
- Map technologies to processes with investment and O&M parameters taken from our technology registry; write investment candidates as TIMES *NCAP* variables when appropriate.
- Convert emissions to commodity emissions and cap trajectories consistent with our policy cases; cost/price inputs follow TIMES units.
- Export either CSV (VEDA/Case Builder style) or GDX via an intermediate table bundle; all conversions and losses are documented in the adapter manifest.

TIMES-NZ → Canonical (Reader)

- Import marginal values on electricity/biomass/emission constraints and translate them into SignalsPack trajectories (*elec_price*, *biomass_price*, *elec_CO_2*, *biomass_CO_2*), taking care to map perfect-foresight period prices to our decision-epoch use (with smoothing or bands).
- Import technology activity and capacity paths for alignment reports in OutlookPack (not dictating our site decisions but providing a comparator portfolio).
- If TIMES provides shadow prices on seasonal timeslices only, the adapter distributes them over our representative or hourly basis using calibrated profiles (declared and tested in the manifest).

Known differences and mitigations. TIMES uses multi-year periods, perfect foresight, and aggregated timeslices; our decision environment uses hourly or representative time and sequential adaptation. The adapter therefore (i) treats imported marginal values as outlook bands rather than precise tariffs, (ii) uses scenario alignment for storylines (policy and price paths) rather than strict co-optimisation, and (iii) reports divergences by boundary definitions rather than tuning parameters.

The same adapter pattern extends to OSeMOSYS/MESSAGEix (process/commodity networks), Spine Toolbox (flexible energy objects), or geospatial feeders (GIS/GeoParquet). To minimise N×M mappings, a small canonical ontology is maintained: controlled vocabularies (technology/fuel/carrier), ISO region codes, units registry, and JSON Schema for artefacts. Where appropriate, Frictionless Data Package descriptors and GeoParquet geometries are attached to artefacts to ease ingestion by third-party tools.

Adapters are first-class code with their own tests:

- **Round-trip tests:** write → read → compare on synthetic cases; tolerance bands are declared (energy ≤1–2%, peaks ≤5–10%, sign and unit checks).
- **Cross-tool invariants:** energy/mass conservation and emissions accounting equalities must hold after mapping.
- **Versioning:** each adapter pins external tool versions and records mapping tables, seeds, and compression settings in the run ledger.
- **Security:** when sharing with external parties, adapters export minimised artefacts (aggregated, non-identifiable) and honour licence/consent fields.

Limitations arise from non-isomorphic temporal structures (hourly vs timeslices), differing semantics of duals (nodal vs system shadow prices), and boundary choices (heat definition, losses). These are

managed by explicit manifests, acceptance diagnostics, and, where signals drive decisions, targeted high-fidelity verification before reporting.

6.8.3 Risks, assumptions, and limitations register

This section formalises the discipline by which modelling assumptions are recorded and justified, risks are monitored and mitigated, and limitations are declared and managed. The register is part of the governance layer: it accompanies every run via the manifest and is version-controlled alongside artefacts. Its purpose is threefold: (i) to make hidden choices explicit; (ii) to enable proportionate mitigation and targeted data collection (e.g., via VoI/VoL); and (iii) to prevent over-interpretation of results beyond the environment's declared scope.

The register is a living, machine-readable catalogue (YAML/TOML/CSV) that lists entries with: ID, class (assumption/risk/limitation), description, rationale, source, quality score, expected effect (directional bias), sensitivity tag, VoI/VoL rating (if applicable), owner, review cadence, date last reviewed, and status (open/mitigated/retired). Each result in OutlookPack references the manifest ID that points to the exact register snapshot used.

Assumptions are unavoidable; the requirement is to state them, defend them, and bound their influence.

Structural assumptions (model form)

Examples: linearised COP/efficiency curves; representative-time clustering with declared weights; Gauss–Seidel (asynchronous) coordination rather than full co-optimisation. Risk if violated: biased duals or masked shortages. Policy: apply §6.9 compression acceptance tests; perform targeted full-hourly spot checks where signals drive investment.

Behavioural/policy assumptions

Examples: ETS price bands as exogenous trajectories; consent lead times represented as minimum commissioning lags; tariff structures (TOU/peak prices) treated as given. Risk: misaligned incentives. Policy: scenario bands and signpost-based adaptation in the exploratory layer.

Technical assumptions (data/parameters)

Examples: boiler efficiencies and header losses from literature where site data are missing; biomass moisture/quality distribution; GXP headroom envelopes. Risk: systematic cost/emission bias. Policy: uncertainty ranges in the Uncertainty Register; VoI score to prioritise data collection (e.g., metered hourly demand, updated GXP capacities).

Example entry (concise):

A-06 | Representative-day set (k-medoids, 24 snapshots per season) | Source: internal clustering pipeline | Quality: B | Expected effect: may understate scarcity peaks | Sensitivity: high for tariff-driven electrification | Review: per epoch or when tariff structure changes | VoI: medium.

Risks are uncertain events or conditions that could compromise validity, tractability, or interpretability. They are tracked with likelihood, impact, triggers/signposts, mitigation, and contingency.

- **Methodological/identifiability risk:** Distinct parameter combinations fit the same outputs (equifinality). Mitigation: anchor with physical constraints; cross-validate against independent series; use scenario discovery to expose degeneracy; report ranges rather than points.
- **Compression bias risk:** Representative-time snapshots distort peaks/ramp scarcity. Mitigation: enforce §6.9 acceptance tests; pin extremes; run local full-hourly verification where duals inform decisions.
- **Adapter/interface risk:** Mis-mapping to external tools (PyPSA/TIMES-NZ) causes unit or semantic drift. Mitigation: adapter round-trip tests; unit registry; signed manifests; small synthetic systems with known optima for smoke tests.
- **Data adequacy/confidentiality risk:** Proprietary site data unavailable or partially redacted. Mitigation: use defensible defaults with wide bands; label quality; VoI to justify data requests; publish only non-identifiable aggregates.
- **Computational risk:** MILPs time out; ensemble growth creates scenario explosion. Mitigation: tiered fidelity; warm-starts; presolve; integer relaxation for screening; DOE-driven scenario reduction; asynchronous re-solve of only affected sites.
- **Policy drift risk:** Rapid ETS/tariff rule changes invalidate outlooks. Mitigation: maintain policy scenario registry; DAPP signposts; rerun with updated bands; document vintage in the Assumption Register.
- **Misuse risk:** Results presented as single “forecast” rather than robust options. Mitigation: default to ensemble/robustness reporting; include regret/satisficing metrics; watermark figures with manifest IDs and scope notes.

Example entry (concise):

R-12 | Duals under compression used for investment trigger | Likelihood: medium | Impact: high | Trigger: Δ(LMP) across iterations > threshold while acceptance tests marginal | Mitigation: full-hourly

spot check for affected weeks; increase cluster count; apply damping | Contingency: defer trigger to next epoch.

Limitations are deliberate scope boundaries at this stage of development; they protect tractability and clarify interpretation.

- **Agent behaviour and market formation**: No endogenous strategic bidding or firm-level adaptive expectations; tariffs and ETS are exogenous storylines.
- **Macroeconomic feedbacks:** No endogenous price–demand feedback from general equilibrium; economy-wide effects are imported via scenarios.
- **Electro-technical detail:** Power-system physics beyond what is required for prices/headroom (reactive power, protection constraints) are modelled externally (e.g., PyPSA) and soft-linked.
- **Process thermodynamics:** Temperature/pressure feasibility enforced via headers and envelopes; full pinch analysis or detailed heat-exchanger networks are out of scope for now.
- **Temporal aggregation:** Representative-time compression is permitted subject to §6.9 tests; commissioning/retirement windows and tariff-critical periods are modelled at full hourly resolution.
- **Geospatial resolution:** GXPs and key feeders are explicit; intra-site piping geography is abstracted unless it binds decisions.
- **Uncertainty semantics:** Deep uncertainty addressed through RDM/DAPP and scenario discovery; probabilities, where used, are illustrative rather than asserted.

Each limitation is paired with an upgrade path (what would be required to relax it) and an impact note (how conclusions should be interpreted given the constraint).

The register is reviewed on a declared cadence (e.g., per decision epoch or upon major data/policy updates). Signposts—measurable indicators such as realised LMP bands, observed GXP headroom utilisation, biomass delivered cost, or ETS auction outcomes—trigger reclassification of entries (e.g., from assumption to observed fact) and re-running of key scenarios. Revisions increment semantic versions; diffs are logged; deprecated entries are archived rather than overwritten to preserve lineage.

Relationship to artefacts and reporting

- **Assumptions → manifests.** Every assumption referenced by a run is cited in the manifest (ID list) and echoed in OutlookPack metadata.

- **Risks → dashboards.** Open high-impact risks surface on a governance dashboard with owner and next action.
- **Limitations → figures/tables.** Each figure or table in the thesis or reports carries a short "scope note" linking to the relevant limitations.

By institutionalising this register, the environment moves from implicit, ad-hoc choices to explicit, reviewable commitments. That shift underpins trustworthy, policy-facing analysis: readers can see not only what was concluded but also which assumptions and risks make those conclusions conditional, and how to update them as evidence accumulates.

6.9 Extensibility roadmap and transition to the proof-of-concept

This chapter has argued for a thin-waist, artefact-governed modelling environment designed to remain interpretable and auditable under sparse, heterogeneous industrial data. The design intent is explicitly progressive, prioritises stable interfaces and disciplined provenance so that fidelity and scope can increase over time without disrupting comparability or breaking the transformation chain from assumptions to results. However, the thesis does not attempt to implement the full long-horizon roadmap here. Instead, it adopts a deliberately bounded proof-of-concept (PoC) as the next step, to demonstrate that the architectural claims made in this chapter can be operationalised end-to-end with a minimal coupling contract and governed reporting.

Accordingly, the near-term roadmap implemented in this thesis is limited to establishing (i) a stable artefact contract sufficient for an end-to-end run, (ii) a deterministic site core that produces interpretable hourly signals from sparse annual evidence, and (iii) a regional evaluation layer that consumes a narrow-exported signal to produce decision-relevant robustness diagnostics. The PoC therefore focuses on one representative industrial case, Fonterra Edendale within the Southland context, using a small number of decision epochs and a controlled uncertainty ensemble, not to provide definitive planning outcomes, but to validate feasibility, traceability, and decision relevance of the modular workflow.

Chapter 7 translates this roadmap into an implementable methodology. It specifies the PoC scope and its acceptance gates; defines the canonical artefacts (e.g., DemandPack and the exported coupling signal); and formalises the bounded uncertainty propagation used in later chapters (screening and paired-futures overlays) while maintaining the thesis's interpretability and provenance requirements. Chapter 8 then reports the resulting artefacts and uses them to assess, within the PoC boundaries, what the environment demonstrably supports and what remains reserved for future extensions. The

broader extensibility roadmap (richer coupling, calibration, multi-site scaling, and cross-carrier expansion) is consolidated in the Conclusions and Perspectives chapter, where it can be framed explicitly as progressive capability growth rather than as a requirement of the present PoC.

Chapter 7
Proof-of-Concept Case Study: Edendale & Southland

7.1 Introduction

This chapter presents the proof-of-concept (PoC) instantiation of the modular site-to-region modelling framework developed in the preceding chapters. Chapter 6 specified the architectural contract; the system boundary, the temporal scaffold of decision epochs, the module interfaces, and the thin-waist artefacts, most notably DemandPack and SignalsPack, that stabilise information exchange across tiers. Chapter 7 implements that contract for a concrete case study: Fonterra's Edendale dairy processing site situated within the Southland regional context.

The purpose of the PoC is methodological rather than predictive. It does not claim a fully calibrated planning model for Edendale, Southland, or New Zealand, nor does it seek to reproduce detailed plant scheduling, steam-network physics, or electricity market operations. Instead, it provides transparent evidence that the proposed architecture operates end-to-end with auditable artefacts and explicit validation gates. Specifically, it demonstrates that heterogeneous case-study inputs can be transformed into standardised, versionable data products; that site-level operations can be evaluated on an hourly chronology; that the site can export a well-defined coupling signal at the regional boundary (incremental electricity demand attributable to electrified heat supply); and that candidate strategies can be stress-tested under deep uncertainty using regret- and satisficing-based robustness measures. In short, the central claim demonstrated in this chapter is that site-level decarbonisation decisions can be evaluated coherently under deep uncertainty while accounting for exogenous regional constraints, without requiring tightly coupled, monolithic optimisation.

Consistent with this aim, the PoC is framed as a one-pass coupling demonstration: DemandPack → proportional dispatch → incremental electricity export → Grid Exit Point (GXP) interface screening and Robust Decision Making (RDM) → robustness summaries. The single-pass structure is a deliberate design choice, ensuring a clear separation of concerns between layers, deterministic and testable outputs, and a tractable workflow suitable for thesis reporting. It also provides a stable methodological baseline against which more sophisticated iterative coordination (e.g., multi-pass signal updating) can be introduced in future work without breaking interfaces.

Several design choices are deliberate to keep the PoC defensible and transparent. The regional electricity layer is represented at the interface level through SignalsPack-style tariff and headroom signals, rather than as a full network expansion model, and upgrade costs are treated as screening-level approximations informed by available planning evidence. Likewise, the PoC emphasises structural validity, time integrity, unit consistency, feasibility and energy closure, and provenance over fine-grained numerical realism. This is consistent with the thesis objective, aiming to show that, a modular architecture can begin from limited, imperfect information while remaining coherent, auditable, and incrementally improvable without breaking interfaces or downstream consumers.

The remainder of the chapter is organised as follows. Section 7.2 justifies the Edendale–Southland case study as a policy-relevant and technically demanding testbed for the site-to-region signalling concept. Section 7.3 describes the case study context and data foundations, including the Edendale site overview, the Southland regional electricity context, and the principal data sources and pre-processing assumptions. Section 7.4 specifies the PoC design choices, including the decision framing across epochs, the modelling scope and boundary, and the artefact conventions adopted to ensure provenance and reproducibility. Section 7.5 presents the site-layer methodology, synthetic hourly heat-demand generation (DemandPack), proportional dispatch with cost and emissions accounting, and the exported incremental electricity signal. Section 7.6 formalises the regional screening and robustness methods, including paired-futures RDM screening and the post-processing robustness overlay used to compare pathway performance under stylised uncertainty. Finally, Section 7.7 documents Integration and reporting gates, and provides the handover to the results and discussion in Chapter 8.

7.2 Case study justification

Edendale and the wider Southland region are selected as the PoC application because they sit at the intersection of (i) high-emitting industrial process-heat demand, (ii) near-term decarbonisation investments already underway, and (iii) binding regional infrastructure and resource constraints. This combination creates a policy-relevant and technically demanding setting in which the proposed modular, uncertainty-aware site-to-region framework can be tested end-to-end, and site-level operational choices materially change electricity requirements at the regional interface, while regional limits (hosting capacity, upgrade options, and fuel supply constraints), condition the feasibility and robustness of site pathways.

From a sectoral perspective, Edendale is one of Fonterra's largest manufacturing sites and a major milk-processing plant with multiple dryers and associated utility systems [41]. Dairy processing is

among the largest contributors to New Zealand's industrial process-heat fossil-fuel use, with coal remaining a dominant fuel in several large South Island facilities [373]. Regional Energy Transition Accelerator (RETA) assessments further indicate that South Island process-heat sites account for a substantial share of New Zealand's direct coal use for energy purposes, with thermal capacity and emissions concentrated in a relatively small number of large facilities [44]. Within this portfolio, Southland's major dairy plants, including Edendale, are therefore disproportionately influential for regional fuel use and emissions trajectories.

Policy and investment dynamics reinforce the relevance of this case. Fonterra's climate roadmap includes commitments to exit coal at remaining coal-fired sites by 2037 and to progress towards net-zero emissions by 2050, supported by interim targets [67]. Edendale is explicitly positioned as a flagship site within this transition, including the commissioning of electrode boilers and staged replacement of coal units [41], [374]. These investments are partly enabled through national policy instruments and co-funding mechanisms (e.g., EECA's GIDI programme) and align with broader objectives to phase out coal for low- and medium-temperature process heat by 2037 [48].

At the regional scale, Southland Murihiku is frequently identified as both highly exposed to coal phase-out and central to achieving national decarbonisation goals. Coal use for process heat has been historically reinforced by the absence of a reticulated natural-gas network and by local lignite availability, while electrification pathways depend on the availability of grid hosting capacity and timely network investment [48]. Regional strategies emphasise that coordinated action across electricity infrastructure, fuel supply chains, and policy incentives is required, particularly because a small number of large sites account for the bulk of regional industrial emissions [44].

Critically, the technical challenge at Edendale extends beyond boiler substitution. Large-scale electrification changes the site's electricity import profile and increases exposure to electricity price volatility, network tariffs, and hosting-capacity (headroom) constraints at the Grid Exit Point (GXP) interface [375]. Biomass-based pathways introduce a different uncertainty set related to fuel availability, logistics, sustainability constraints, and competition for feedstock [44]. Planned production expansions may further increase heat demand, strengthening the coupling between site operations and regional infrastructure requirements [376]. Collectively, these characteristics align directly with the modelling requirements established earlier: high site-level complexity and staged investments, embedded within a constrained regional context where shared limits and uncertain trajectories shape feasible transition pathways.

Finally, this case offers decision relevance beyond the PoC demonstration and provides a direct test of the framework's core design claims: thin-waist artefacts can carry consistent, auditable signals across

tiers, while robustness screening can evaluate pathway performance. The visibility of Edendale's transition in national decarbonisation discourse, together with Southland's strategic role in New Zealand's export economy, means that the methodological insights and robustness comparisons developed here are likely to be informative for agencies (e.g., EECA, MBIE), network operators, and other major process-heat users. Accordingly, the PoC is used both as a methodological proof and as a credible exemplar for how robust, least-regret industrial transition strategies could be evaluated under deep uncertainty in the New Zealand context.

7.3 Case study context and data overview

To instantiate the framework contract defined in Chapter 6, the Edendale–Southland PoC requires a case description that can be expressed in the thin-waist artefacts used for coupling and evaluation. Accordingly, this section summarises (i) the site attributes required to construct DemandPack and parameterise the site dispatch; (ii) the regional characteristics that motivate a binding SignalsPack electricity interface and a screening-level upgrade menu; and (iii) the datasets used to parameterise the PoC, distinguishing reported values from stylised assumptions. The emphasis therefore on PoC-relevant abstractions, annual energy magnitudes, epoch-tagged utility portfolios, availability logic, and a GXP-level hosting-capacity envelope, rather than full plant reconstruction.

7.3.1 Edendale site overview

Edendale is Fonterra's southernmost manufacturing facility and one of its largest and most complex multi-product sites. It has historically been described as New Zealand's largest raw-milk processing plant, with capacity rising to around 15 million litres of milk per day after successive expansions [377]. The site comprises multiple production trains (including four milk-powder spray dryers) and additional plants for cheese, cream and fat-derived products, supported by utilities such as refrigeration and wastewater treatment. For the PoC, Edendale is modelled as heat-and-electricity demand node served by a portfolio of utility units rather than as a process-resolved production model. This choice reflects the current data reality: while the site's major assets and transition investments are publicly documented, process-level schedules and measured steam balances are not. Accordingly, hourly heat demand is generated synthetically from annual totals with stylised seasonality and diurnal shaping and is then dispatched across the available utility units under feasibility and operating-logic constraints.

Committed electrification investments provide the principal structural change across PoC epochs. The first large electrode boiler (around 20–30 MW) commissioned in the mid-2020s is reported to reduce site emissions by approximately 47,000–47,500 tCO_2e per year (about 20% of site emissions) [41]. Two

additional electrode boilers (each 30 MW) are reported for commissioning by 2027, replacing further coal capacity and enabling additional steam supply for new production investment [67], [377]. The PoC treats the near-term electrification investments as committed, exogenous changes to the utility portfolio (one electrode boiler commissioned in 2024; two further units by 2027), and defines the key decision lever for subsequent analysis as the replacement of the remaining coal capacity via (i) additional electrification or (ii) a biomass pathway. The purpose here is not to reconstruct historical dispatch but to provide a transparent, auditable instantiation of the site module that can generate the thin-waist outputs required for coupling (in particular, incremental electricity demand associated with electrified heat supply).

Planned production expansion introduces additional demand uncertainty. The UHT cream plant announced as an investment of approximately NZD 150 million is expected to increase production volumes and heat demand [376]. Detailed process-resolved production schedules are not available; therefore, the PoC represents growth through demand multipliers and stylised reshaping parameters within the synthetic demand generator, rather than explicit process scheduling. This choice preserves interface stability while allowing later refinement if process-level data become available.

Electrification at Edendale couples directly to the regional electricity interface. The site is supplied through the Edendale grid exit point (GXP), which also serves local rural load. RETA assessments indicate that spare capacity at this interface is limited and that further electrification may require significant reinforcement [375]. In the PoC, these constraints are represented in reduced form through (i) an hourly GXP headroom envelope and (ii) a tariff signal, both provided via SignalsPack. These are instantiated as time-aligned hourly traces at the interface level, parameterised from available planning evidence and explicitly treated as screening-level inputs. The resulting coupling variable is the incremental electricity demand attributable to electrified heat supply, exported from site dispatch as an hourly series for evaluation against headroom and upgrade strategies in later sections.

7.3.2 Southland regional context

Southland Murihiku is a primary-production region whose economy is shaped by dairy processing, meat works, forestry and associated logistics. These activities produce a disproportionately large process-heat footprint relative to population. Regional emissions reporting indicates that stationary energy contributes a substantial share of gross regional emissions and that manufacturing and commercial activities dominate stationary-energy scope-1 emissions [378]. This makes Southland a suitable context for examining industrial heat decarbonisation under regionally binding constraints.

The regional decarbonisation agenda focuses strongly on stationary heat and transport fuels as the two largest abatement levers. The Southland Murihiku Regional Energy Strategy 2022–2050 sets an explicit vision for clean, resilient and affordable energy and includes goals for decarbonising commercial and industrial thermal heat. This regional framing aligns with national policy direction, including the planned phase-out of coal for low- to medium-temperature process heat by 2037 and economy-wide net-zero targets by 2050 [48]. For the PoC, these policy signals motivate the "committed pathway" framing and define the decision relevance of timing, feasibility and least-regret infrastructure choices.

Southland's relevance to the PoC lies less in regional averaging and more in the fact that it imposes binding shared constraints on industrial heat transition: electricity network hosting capacity and the availability/cost of low-carbon fuels. RETA frames electricity availability as a two-stage issue: adequacy at the wholesale level and deliverability through transmission/distribution networks to GXPs and sites, where the practical constraints are (i) spare capacity, (ii) upgrade costs, and (iii) the level of connection security required [44]. This is precisely the interface the PoC needs: Edendale's electrification decisions translate into incremental electricity demand that must be screened against local hosting limits and upgrade choices.

RETA's network assessment identifies a subset of "major complexity" industrial connections in Southland, explicitly including Fonterra Edendale, where electrification requires material investment in Transpower assets (substations/lines) and where delivery timelines are on the order of two to four years before consents and easements are considered [375]. For Edendale, RETA provides a structured menu of connection options (existing load, additional electrified demand, security level, and indicative capital cost) that is suitable for screening-level modelling and strategy comparison:

- Option 1: Existing 30 MW + Additional 21 MW (N security), $9.60m
- Option 2: Existing 30 MW + Additional 32 MW (N-1 security), $34.10m
- Option 3: Existing 30 MW + Additional 97 MW (N security), $55.90m [44].

In the PoC, these RETA options provide the empirical backbone of the regional-interface upgrade menu, specifying indicative capacity increments and associated screening-level capital costs for the electricity hosting-constraint assessment. They are treated explicitly as order-of-magnitude inputs suitable for comparative robustness screening rather than detailed network planning. To ensure the menu spans the full electrification envelope implied by replacing the remaining coal capacity, an additional high-capacity option is included as a PoC extension; this option is explicitly flagged as an

assumption and is used only for sensitivity and robustness screening, with implications discussed alongside the PoC limitations.

Biomass is the second binding regional factor. RETA estimates that, averaged over the next 15 years, Southland has around 380,000 t (2,740 TJ) per annum of woody biomass that could be recovered and/or diverted to bioenergy markets in the near term—sufficient for a "pragmatic" conversion pathway but insufficient for a high-demand scenario in which all large sites convert without drawing on higher-grade logs. The same assessment provides delivered cost ranges (including procurement, processing, storage, and delivery): $9–12/GJ for residues, $14–18/GJ for diverted export chip/low-grade logs, and up to $21–23/GJ (unpruned logs) and $29/GJ (pruned logs) [44]. For the PoC, biomass therefore enters not as a deterministic "solution", but as a scenario environment (e.g., favourable vs constrained supply/cost conditions) that shifts the relative attractiveness and robustness of electric boiler (EB) vs biomass boiler (BB) pathways under uncertainty.

Taken together, Southland provides the necessary regional conditions for a meaningful site↔regional PoC: electrification creates non-trivial hosting-capacity and lead-time constraints with a tractable upgrade menu, while biomass availability and delivered costs are substantial yet uncertain. RETA also explicitly notes that the upgrade costs reported are indicative and appropriate for screening analysis, which matches the PoC's intended level of regional representation [44].

7.3.3 Data sources

The PoC draws on a deliberately small set of high-leverage sources that map directly onto the Chapter 6 artefacts:

1. Site transition commitments and asset parameters (public Fonterra/EECA/GIDI disclosures) used to fix the committed 2024/2027 utility additions and to bound plausible post-2027 options for the remaining coal capacity [15], [41], [44].
2. Regional electricity constraints and upgrade options (Southland RETA network assessment) used to parameterise SignalsPack components (tariff and headroom signals and an upgrade menu) consistent with screening-level RETA evidence [375].
3. Regional biomass availability and delivered cost ranges (Southland RETA biomass assessment) used to define the biomass scenario environments and fuel cost parameters for the biomass pathway comparison [44].

4. Demand and climate drivers (RHDD/MBIE/NIWA as applicable) used to anchor annual energy magnitudes and construct seasonal/diurnal shaping assumptions for DemandPack (with documented parameters where empirical gaps remain) [7], [16], [379].

All inputs are transformed into versionable, time-aligned artefacts (DemandPack and SignalsPack) using reproducible preprocessing scripts and validation gates (timestamp integrity, unit consistency, and annual energy closure). Where values are screening-level (notably network upgrade costs and biomass supply/cost envelopes), this is treated as an explicit modelling assumption consistent with RETA's stated intent.

7.4 PoC Design: decision framing, scope, and artefact discipline

This section specifies the design commitments that make the PoC executable, comparable across epochs, and auditable. It formalises (i) the decision framing and epoch structure used to represent staged transition pathways; (ii) the system boundary and the rationale for the one-pass coupling design; and (iii) the artefact conventions, run organisation, and acceptance gates used to ensure traceability, unit consistency, and fair comparison under uncertainty. Throughout, the emphasis is on disciplined transformation from heterogeneous case inputs into standardised, versionable artefacts that can be consumed consistently by downstream modules.

7.4.1 Decision framing and epochs

The PoC is defined on a discrete set of epoch tags, each representing a self-contained planning snapshot with an epoch-specific technology portfolio, parameters, and signals. An epoch tag therefore functions as a *minimal reproducible specification* of the site–region evaluation problem. The annual heat requirement is converted into an hourly DemandPack profile, dispatched across an epoch-specific utility set, and, where electrification is present, translated into an exported incremental electricity time series that is screened at the grid interface. All epochs are evaluated on a common hourly chronology (UTC), enabling direct comparison of both time-series behaviour and annual performance summaries.

The PoC uses five epoch tags:

- **2020:** Baseline configuration.
- **2025:** Near-term transition snapshot.
- **2028:** Mid-term transition snapshot (post-commissioning of committed assets).
- **2035_EB:** 2035 electrification pathway (electric boiler–dominant).

- **2035_BB:** 2035 biomass pathway (biomass boiler–dominant, with electrification limited to the committed prefix).

The epoch structure distinguishes between a committed prefix and a strategic fork. The committed prefix treats announced near-term electrification investments as exogenous, shared changes to the site's utility portfolio across the intermediate epochs (2025 and 2028), ensuring both long-term pathways inherit a common transition history up to the fork point. The 2035 fork then represents the primary strategic decision examined in the PoC, which is the replacement of the remaining coal-derived heat supply by 2035 via either further electrification (EB) or a biomass pathway (BB). This framing creates a controlled comparison in which the two 2035 pathways differ only in the post-2028 replacement choice, reducing confounding and enabling straightforward attribution of differences in outcomes.

Backcasting from the required end state fixes the decision. National policy plans the phase-out of coal for low- and medium-temperature process heat by 2037, and Fonterra has committed to exit coal by 2037 [44], [63]. With the committed electrode boilers treated as given, the remaining decision is how the last coal-derived heat supply is replaced by 2035. It is taken from the site operator's perspective, but its consequences may extend to the Edendale Grid Exit Point. Table 7.1 states the problem in the XLRM form used in robust decision making (Chapter 5).

Table 7-1 XLRM framing of the proof-of-concept decision

Element	Proof-of-concept specification
Levers (L)	Site pathway: 2035_EB or 2035_BB. Grid response: upgrade option S0–S4, or the least-cost option per future (AUTO).
Uncertainties (X)	Grid screen: headroom (U_head), incremental load (U_ΔE), upgrade capital cost (U_capex), value of lost load (U_VOLL), consenting uplift (U_cons). Pathway overlay: electricity, biomass and ETS price multipliers.
Relationships (R)	DemandPack → proportional dispatch with hourly heat balance → exported ΔE → GXP headroom screen with upgrade menu → site, grid and system cost ledgers
Measures (M)	Annual site cost; direct CO_2; selected upgrade and annualised cost; annual unserved energy; system cost; regret; satisficing rate; win rate

To support consistent reporting and downstream robustness screening, each epoch produces a common set of indicators derived from the dispatch summary (site layer) and, for 2035, the RDM

screening summary (regional interface layer). These indicators are defined here to establish what is measured and how it is recorded.

Site-layer indicators (dispatch summary; TOTAL row):

- **Cost accounting:** *annual_total_cost_nzd*, with components such as *annual_fuel_cost_nzd*, *annual_electricity_cost_nzd*, and *annual_carbon_cost_nzd*.
- **Electricity exposure:** *annual_electricity_MWh* and *avg_electricity_tariff_nzd_per_MWh*. Where electricity accounting may appear indirectly through unit-level fuel accounting for electrode boilers, the PoC reports effective totals (e.g., *annual_electricity_MWh_effective*, *annual_electricity_cost_nzd_effective*) to ensure EB–BB comparisons remain consistent under alternative accounting representations.
- **Emissions accounting:** *annual_co2_tonnes* and, where enabled, the associated carbon-cost proxy (*annual_carbon_cost_nzd*).
- **Feasibility and penalties:** *annual_unserved_MWh* and associated cost fields, used as feasibility diagnostics and as stress indicators for binding constraints or insufficient capacity.

Regional-interface indicators (screening/RDM; 2035 only):

- Upgrade choice and magnitude: *selected_upgrade_name*, *selected_capacity_MW*.
- Annualised upgrade cost: *annualised_upgrade_cost_nzd*.
- Residual exceedance/shedding: *annual_shed_MWh*, *shed_fraction*, and *unserved_peak_MW*.
- Total interface cost: *total_cost_nzd* (annualised upgrade cost plus shed cost), used as the principal comparative measure in the screening layer.

Together, these indicators form the common reporting basis for within-2035 pathway comparison (EB vs BB) and for cross-epoch summaries (e.g., baseline → mid-term → long-term), while preserving clear separation between site-layer operational evaluation and regional-interface feasibility screening.

7.4.2 System boundary and coupling rationale

The PoC adopts an explicit system boundary aligned with the framework contract in Chapter 6, separating within-fence site operations from regional feasibility conditions at the point of grid connection. The boundary is defined to be sufficiently rich to demonstrate end-to-end signal flow and artefact discipline yet sufficiently constrained to remain auditable and defensible at PoC maturity. In

particular, the PoC is designed to demonstrate how site decisions generate interface-relevant signals and how those signals can be screened under uncertainty, without implying that the PoC performs integrated co-optimisation of site and grid systems.

Site layer (within-fence operational evaluation)

Edendale is represented as an aggregate heat demander served by an epoch-tagged portfolio of utility units (e.g., coal boilers, electrode boilers, biomass boilers) subject to simplified operational logic and availability constraints. Within this boundary, the PoC includes: (i) synthetic hourly heat demand (DemandPack) derived from annual totals and shaping assumptions; (ii) unit-level heat supply allocation via a deterministic dispatch kernel (with alternative kernels implemented but reported results use proportional dispatch); and (iii) accounting ledgers for fuel use, electricity use, direct combustion emissions, and cost components (fuel, electricity, ETS-style carbon proxy, and feasibility penalties where applicable). The site layer does not attempt to reproduce detailed steam-network thermodynamics, multi-pressure headers, condensate return constraints, or production scheduling; these are deliberately abstracted into a single heat-demand stream and unit conversion efficiencies to preserve transparency, closure, and interpretability.

Regional layer (interface-level proxy screening)

The regional representation is implemented as an interface-level proxy rather than a full electricity system or network expansion model. In the PoC, the regional proxy comprises: (i) an hourly hosting-capacity envelope at the relevant GXP (headroom); and (ii) a discrete upgrade menu with indicative capacity increments and annualised costs, evaluated under uncertainty, including potential load shedding valued at a Value of Lost Load (VOLL) parameter for screening. This is sufficient to assess whether a candidate electrification pathway is likely to be feasible at the point of connection and to characterise reinforcement exposure in a transparent, inspection-ready manner. The PoC does not represent wholesale market clearing, nodal price formation, generator unit commitment, or explicit transmission/distribution topology; these drivers are conservatively collapsed into interface signals and explored through uncertainty multipliers in the robustness layer.

Importantly, the regional proxy is framed as a replaceable module that produces or consumes the same thin-waist artefacts. In later work, the proxy can be substituted by a regional electricity module (e.g., PyPSA-based dispatch and constraint representation) while preserving the same SignalsPack-style interface (time-indexed tariffs, hosting constraints, headroom, and potentially marginal emissions factors). This substitution changes fidelity, not architecture, the site layer remains insulated from the internal formulation of the regional layer by the artefact contract.

Boundary register

Applying the decision-first boundary principle (Section 6.3.3), Table 7.2 records which consequences of the 2035 choice enter the comparison, at what fidelity, why, and how each excluded consequence could enter later.

Table 7-2 Boundary register for the 2035 decision

Consequence	Status	Reason	Later inclusion
Site heat balance, energy use, cost, CO_2	Included	Drives site cost and emissions	Richer dispatch kernels (§7.5.3)
Grid headroom, reinforcement, unserved energy	Included (proxy)	Differs between pathways	Power-flow or PyPSA module, same ΔE interface
Biomass regional supply	Proxied by cost; relevance unresolved	No regional supply data	Supply module via SignalsPack
Demand timing and peaks	Proxied (synthetic)	No metered site data (§7.3.1)	Metered weeks, same DemandPack
Heat pumps and hybrids	Excluded	Decision is replacing the remaining coal steam capacity	Added unit with COP curve (§6.7.3)

Two features of this register shape interpretation. First, the boundary is asymmetric: the electrification pathway carries an explicitly modelled grid-interface consequence, whereas the regional consequences of the biomass pathway enter only through its delivered cost. The results can therefore show that a regional consequence of electrification is visible and evaluable; they cannot show that the biomass pathway is regionally unproblematic. Second, every exclusion carries a reason and an entry route, so the comparison can be extended where the diagnostics show it matters, rather than enlarged in advance.

Signals discipline and directionality

To keep causality and provenance unambiguous, the PoC separates signals provided to the site, the signal exported from the site, and the regional evaluation of that export. In the reported PoC execution:

- Signals consumed by the site (downward-flow context): Site dispatch consumes exogenous context signals used for accounting and, in non-reported dispatch modes, for cost-based allocation logic. These include electricity tariffs and, where enabled, an ETS-style carbon-price proxy applied to combustion emissions. These inputs may be flat or time-varying series; in all

cases they are treated as externally specified within a run and recorded as part of the run provenance.

- Signal produced by the site (upward-flow export): The coupling export is the incremental electricity demand attributable to electrified heat supply, written as an hourly series aligned to the common UTC chronology. This exported series is a thin-waist artefact designed to be consumed by the regional layer (and other downstream analyses) without requiring direct access to site-internal model state.
- Regional proxy screening (evaluation without feedback in the PoC): The regional screen consumes the incremental export together with headroom and the upgrade menu to compute exceedance, potential shedding exposure (via VOLL), and indicative upgrade exposure under uncertainty. These outputs are recorded as auditable artefacts for robustness analysis and reporting; they do not alter site dispatch quantities or update tariffs in the reported PoC execution.

This directionality is a methodological choice for the PoC reporting boundary, not a limitation of the framework concept. The key point is that the PoC demonstrates a disciplined signal pathway and artefact integrity first, providing a reliable base for extensions.

Coupling choice and rationale

The coupling between tiers is achieved through a single, explicitly defined export from the site layer. Incremental electricity demand attributable to electrified heat supply, expressed as an hourly series (MW) on the same timestamp base as the regional headroom and tariff signals. This design is methodological rather than technical. It is selected because it:

- Preserves auditability through separation of concerns. The site layer can be validated independently (demand integrity, heat balance, carrier accounting, costing and emissions ledgers), and the regional layer can be validated independently (time index integrity, exceedance logic, upgrade menu evaluation, cost closure). This reduces the risk that errors are obscured by feedback interactions.
- Ensures determinism and reproducibility. A single forward pass yields deterministic outputs given fixed inputs, simplifying provenance tracking, regression testing across epochs and code revisions, and fair comparison between pathway variants.
- Matches thesis-appropriate scope. The purpose of the PoC is to establish that thin-waist artefacts can support end-to-end evaluation of site decisions under shared regional feasibility

constraints, including robustness screening under uncertainty, without requiring the additional behavioural assumptions and governance choices that accompany iterative coordination.

Architectural extensibility to multi-pass coordination and co-optimisation

Although the reported PoC is single-pass, the architecture is deliberately engineered so that multi-pass coordination and iterative co-optimisation can be introduced as controlled extensions. The essential requirement for iteration is that the regional layer returns actionable scarcity signals that can alter subsequent site decisions. These scarcity signals can take two complementary forms:

- Price-type coordination signals (marginals and dual-inspired adders). The regional layer returns time-varying scarcity adders (e.g., reflecting binding hosting constraints or system stress) that can be encoded as additional tariff components and applied by the site layer in a subsequent solve.

- Quantity-type coordination signals (caps and allocations). The regional layer returns explicit allocations or limits (e.g., time-varying headroom caps, annual or seasonal biomass allocations, emissions allocations) that the site layer must respect as feasibility constraints in a subsequent solve.

This dual design is valuable, price signals are scalable and consistent with decomposition thinking (particularly in multi-site settings), whereas quantity caps provide conservative coordination when feasibility guarantees are required. Because the interface already separates internal model logic from exchanged variables, introducing iteration does not require redesigning the site or regional formulations. Instead, iteration is implemented by: (i) extending the exchanged SignalsPack fields (to include scarcity prices and/or caps), and (ii) adding an orchestration wrapper that executes sequential passes while recording provenance and convergence diagnostics. The PoC already enforces the preconditions for this extension with timestamp alignment on a common UTC chronology, schema-stable artefacts, and closure checks at each layer.

A practical staged pathway to full coordination is envisaged:

- Two-pass exchange (first-order feedback): The regional module evaluates the site export and returns updated SignalsPack inputs (e.g., tariff adders and/or tighter headroom limits). The site is re-evaluated once using these updated signals to quantify first-order implications for costs, dispatch choice (where cost-aware kernels are used), and incremental export.

- Limited multi-pass exchange (stabilisation and diagnostics): A small number of additional passes (e.g., three) can be used to test whether exchanged signals and exports stabilise, with explicit reporting of pass-to-pass deltas (e.g., changes in peak incremental MW, binding hours, or tariff-weighted electricity costs). This can be framed as a coordination sensitivity rather than a claim of full convergence.

- Gauss–Seidel style iteration (full coordination): The site and regional solves are repeated until a convergence rule is met (e.g., changes in exported incremental electricity and returned scarcity signals fall below tolerances), subject to a maximum iteration limit and optional damping. All passes are logged as separate, versioned artefacts, preserving traceability.

The thesis reporting boundary remains clear, the PoC demonstrates one-pass signalling and robust screening and establishes the artefact and governance preconditions required for later iterative coordination.

7.4.3 Artefact model, run organisation, and acceptance gates

The PoC applies an artefact discipline designed to make results traceable, comparable across epochs, and auditable. This section specifies (i) the canonical artefacts produced by each layer and their role in site to regional coupling; (ii) run-organisation conventions that prevent accidental mixing across epochs, pathway variants, or sensitivity runs; and (iii) acceptance gates that must be satisfied before outputs are admitted to the thesis reporting set. Every KPI reported in Chapter 8 must be reproducible from a named artefact generated under an explicit bundle–epoch–run identifier, with integrity checks recorded alongside the artefacts they validate.

Outputs are written under a bundle namespace with further scoping by epoch tag and dispatch run identifier:

- Bundle (*bundle*): An immutable container for a coherent thesis reporting package generated under a consistent code and configuration state (including input tables and configuration files). Bundles are the unit of reporting provenance: all tables and figures included in the thesis must resolve to artefacts within the same bundle.

- Epoch tag (*e*): The planning snapshot (2020, 2025, 2028, 2035_EB, 2035_BB), which determines the epoch-specific portfolio, parameters, and scenario framing.

- Run-ID (*run*): A dispatch run label (e.g., a named dispatch kernel configuration), allowing controlled within-epoch variants without redefining the epoch. Run-IDs are used for sensitivity

variants or alternative kernels while preserving the same upstream DemandPack and the same reporting boundary.

This naming discipline supports unambiguous referencing of outputs (*bundle/epoch/run*) and supports extensibility (additional epochs, dispatch kernels, or expanded futures) without breaking the reporting contract.

The pipeline is organised as a sequence of layer outputs that can be inspected independently. The core "thin-waist" contracts are:

1. DemandPack (demand artefact). an hourly heat-demand series on a common UTC chronology, with a run manifest recording configuration references and provenance metadata (including demand closure and timestamp integrity checks). DemandPack is the single authorised source of site heat demand for dispatch in the PoC.

2. Site dispatch outputs (operational ledgers). Hourly dispatch records (long and wide forms) and an annual summary table used as the primary KPI source. These outputs provide a transparent ledger for heat balance, carrier use, emissions, and cost components.

3. Exported coupling signal (site to regional interface). An hourly incremental electricity series attributable to electrified heat supply. This export is the sole coupling variable in the reported PoC and is defined at the same time index as DemandPack and interface signals. Where electrified units are absent, the export is either omitted by design or recorded as identically zero, with that case documented in run metadata.

4. Regional proxy outputs (2035 screening and RDM). For the 2035 pathways, the regional proxy consumes the exported incremental electricity series and evaluates it against headroom and a discrete upgrade menu under a canonical futures set, producing screening summaries and comparison tables.

5. Thesis pack (curated reporting set). A curated set of tables and figures with a pointer map that resolves each curated item to a source artefact within the same bundle. This functions as a reporting-oriented analogue of an OutlookPack and does not introduce new model results, but registers and stabilises the artefacts used in the thesis narrative.

A deliberate design commitment is that this artefact set remains small and stable. Each artefact has a clear interpretation, stable schema, and a defined downstream consumer; richness is added through additional artefacts and optional columns, not through ad hoc changes to the core contracts.

Although the reported PoC is single-pass, the artefact model is engineered so that additional passes can be introduced later as a controlled extension without changing the thin-waist contracts. The enabling conditions are already enforced:

- all exchanged quantities are timestamp-aligned hourly series on a common UTC chronology;
- artefacts are schema-stable, versioned, and referenced by canonical paths rather than implicit in-memory state; and
- closure and integrity checks are recorded at the layer that produces each artefact.

Under a future multi-pass coordination scheme, additional pass-indexed artefacts would be introduced (e.g., *pass01*, *pass02*) to store returned scarcity signals (price adders and caps) as SignalsPack extensions, together with convergence diagnostics. This would be an orchestration-layer change, the site and regional solvers would continue to consume and produce the same base artefacts, with additional signal fields introduced in a versioned, backwards-compatible manner.

A central fairness requirement is that the 2035 pathway comparison is conducted under paired futures. EB and BB must be evaluated under identical uncertainty realisations so that differences in outcomes can be attributed to pathway choice rather than sampling variation. This is enforced by requiring a canonical per-bundle *futures.csv* containing a unique *future_id* set; and the same *future_id* set to appear in both pathway screening summaries (2035_EB and 2035_BB). Any deviation (e.g., incomplete future coverage caused by a failed run) is treated as a reporting failure unless explicitly documented, justified, and excluded from comparative claims.

Before a bundle is treated as thesis-ready, the pipeline applies acceptance checks that act as reporting gates. The gates are grouped into three categories:

- Completeness. Required artefacts exist for declared epochs and pathways (including DemandPack, dispatch summaries, exported incremental electricity, and 2035 screening/RDM outputs where applicable).
- Integrity. Time index and units are consistent (hourly chronology, UTC timestamps, no missing or duplicate hours) and closure checks pass (DemandPack annual closure; dispatch heat balance and accounting reconciliation within tolerance).
- Consistency and fairness. EB and BB satisfy the paired-futures contract; electricity accounting is comparable (including the use of “effective/derived” electricity totals where required); and feasibility indicators (unserved energy) are explicit and correctly costed (never silent or implicit).

These gates provide the methodological basis for claiming that Chapter 8 results are derived from a complete and internally consistent execution of the PoC contract. Table 7.3 registers the minimum artefact set and the associated acceptance gates. If a required artefact is missing or fails its acceptance checks, the corresponding run is excluded from the reporting set.

Table 7-3 Minimum PoC artefact set and acceptance gates (compact register)

Layer	Artefact (canonical location)	Acceptance gates (minimum)
Demand	*epoch<e>/demandpack/meta/run_manifest.json*	Manifest present; referenced configs/inputs resolved; closure/chronology checks recorded.
Demand	*epoch<e>/demandpack/demandpack/hourly_heat_demand_<e>.csv*	8760 hourly UTC; no gaps/dupes; finite values; annual closure within tolerance.
Site	*epoch<e>/dispatch/<run>/site_dispatch_<e>_long.csv*	Schema valid; finite values; heat balance holds within tolerance (incl. UNSERVED).
Site	*epoch<e>/dispatch/<run>/site_dispatch_<e>_wide.csv*	Chronology aligned to DemandPack; totals consistent with long-form aggregation.
Site	*epoch<e>/dispatch/<run>/site_dispatch_<e>_summary.csv*	Required KPI fields present; component sums reconcile; unserved explicit (never implicit).
Signal (export)	*epoch<e>/dispatch/<run>/signals/incremental_electricity_MW_<e>.csv*	8760 hourly UTC; aligned chronology; non-negative under stated convention; optional/zero case documented.
Signals (used)	*epoch<e>/dispatch/<run>/signals_used/tariff_nzd_per_MWh_<e>.csv*	8760 hourly UTC; aligned chronology; unit NZD/MWh; matches costing basis used in summary.
Uncertainty	*rdm/futures.csv (per bundle)*	Unique *future_id*; canonical paired set for EB/BB; multipliers present or defaults explicit.
Screen	*rdm/rdm_{summary,matrix}_2035_{EB, BB}.csv*	Required columns; finite costs; consistent strategy IDs; paired futures validated.
Compare	*rdm/rdm_compare_2035_EB_vs_BB.csv and compare_pathways_2035_EB_vs_BB.csv*	Valid joins on *future_id* (paired subset); deltas computed only on paired futures; any exclusions documented and justified.
Reporting	*thesis_pack/figures/, thesis_pack/tables/, thesis_pack/pointers.md*	Pointers resolve within same bundle; stable filenames; no dependence on scratch paths; curated artefacts are read-only derivatives.

7.5 PoC site-layer implementation

This section documents the site-layer instantiation of the framework for the Edendale PoC. The site layer has two responsibilities: (i) transform case-study inputs into a standardised hourly demand artefact (DemandPack) and (ii) evaluate an epoch-tagged utility portfolio against that demand on a common chronology, producing auditable ledgers for heat balance, costs, and emissions, plus the single exported coupling signal required by the regional interface layer. The site layer is therefore designed as an accounting-faithful operational evaluator, capable of producing internally consistent, traceable quantities, rather than a fully optimising plant scheduler.

Consistent with Chapter 6, the PoC applies a thin-waist contract in which DemandPack carries the upward-flow demand series (hourly heat demand in MW, and co-reported electricity demand if required), while SignalsPack provides the downward-flow exogenous context needed for costing and interface alignment (notably electricity tariff inputs and, where enabled, a carbon-price proxy for emissions costing). Grid-interface feasibility signals such as GXP headroom are treated as interface-layer inputs for regional screening in the reported PoC; they do not feed back into the reported site dispatch. Site dispatch then produces: (a) hourly heat supply by unit (including any unserved residual), (b) annual KPI summaries for costs and emissions, and (c) an exported coupling signal, an hourly incremental electricity series attributable to electrified heat supply, aligned to the same UTC timestamp index as the interface inputs.

The site evaluator is deliberately implemented as a kernel-swappable component. The artefact contracts (DemandPack in; dispatch ledgers out; incremental-electricity export out), the validation gates, and the downstream consumers are invariant, while the internal dispatch rule may be substituted without changing the pipeline interfaces. Reported PoC results use a deterministic proportional allocation kernel for transparency and auditability. However, the codebase also includes two additional swap-in kernels, an hourly LP dispatch kernel and a blockwise “optimal subset” kernel, which accept the same input schema and produce the same output ledgers. Chapter 8 results are derived from the proportional kernel only; alternative kernels are documented as implemented, interface-compatible capabilities not exercised in the reported PoC bundle.

7.5.1 DemandPack construction

For each epoch tag e, DemandPack produces a single hourly site heat-demand series on a fixed annual grid, expressed as interval-average thermal power $D_{t,e}^{heat}$ in MW over each hourly interval. In the PoC, DemandPack is the only authorised source of site heat demand for dispatch; its validation gates are therefore treated as a hard precondition for all downstream computations, including dispatch feasibility, costing, incremental electricity export, and subsequent regional screening and RDM evaluation. No downstream module may overwrite or re-timebase demand. This deliberately elevates demand construction from a “data-preparation step” to a contracted artefact whose integrity underpins all reported results.

DemandPack writes an epoch-scoped output file (*hourly_heat_demand_<epoch>.csv*) within the bundle- and epoch-specific DemandPack directory. The workflow supports more than one legacy location pattern for backward compatibility; however, the dispatch module resolves one canonical

input path per run and records the resolved path in run metadata. This prevents ambiguity, avoids accidental mixing across epochs or runs, and preserves provenance for thesis reporting.

The minimum schema required by downstream consumers is:

- *timestamp_utc*: ISO-8601 UTC timestamp (canonicalised to UTC; strictly hourly)
- *heat_demand_MW*: interval-average thermal demand in MW (non-negative).

Optional diagnostic columns may be written (non-contractual), including intermediate shaping-factor components (e.g., seasonal, weekday, diurnal multipliers), scenario labels (epoch and variant tags), and if stochastic perturbations are enabled, noise realisation identifiers and seed tags. This separation between a stable contract schema and optional diagnostics maintains interface stability while permitting incremental richness and auditable review.

Formally, let $t = 1, \dots, T$ with $T = 8760$ for a standard (non-leap) year representation, and define the timestep $\Delta t = 1$ hour. Demand is represented as:

$$D_{t,e}^{heat}[MW] = interval - average\ thermal\ power\ over\ hour\ t,$$

so that annual thermal energy is computed dimensionally correctly as

$$E_e^{ann}[MWh] = \sum_{t \in T} D_{t,e}^{heat} \Delta t \quad .$$

This convention is enforced consistently so that all subsequent energy, emissions, and cost calculations remain well-posed (e.g., $cost = tariff[NZD/MWh] \times energy[MWh]$) and comparable across epochs.

The synthetic profile is constructed from an interpretable, factor-based shape $s_{t,e}$ (dimensionless), derived from a small set of explicit components parameterised via configuration and factor files (e.g., seasonal factors, weekday factors, and hour-of-day profiles):

$$s_{t,e} = f_{m(t),e}^{season} \times f_{d(t),e}^{wk} \times f_{h(t),e}^{hr} \quad , \qquad s_{t,e} \geq 0 \quad ,$$

where $m(t)$, $d(t)$, and $h(t)$ map each hour to its month (season), day-type, and hour-of-day categories. The final demand series is obtained by scaling this shape to an annual energy target E_e^{ann}:

$$D_{t,e}^{heat} = \alpha_e s_{t,e}, \qquad \alpha_e = \frac{E_e^{ann}}{\sum_{t \in T} s_{t,e} \Delta t}.$$

This scaling guarantees annual energy closure by construction, subject only to floating-point tolerance. In the reported PoC bundle, the intra-annual shaping components are held fixed across

epoch tags (i.e., $s_{t,e} = s_{t,2020}$), so the hourly demand *shape* is invariant by construction. Any epoch-to-epoch differences in $D_{t,e}^{heat}$ therefore enter only through the declared annual energy target E_e^{ann} (and any configured peak cap), rather than through inferred production or operational dynamics.

Two optional refinements may be activated via configuration; neither introduces feedback from dispatch outcomes:

- **Temperature sensitivity (optional):** If enabled, DemandPack applies a bounded adjustment based on an external temperature series T_t and reference T^{ref}, followed by re-normalisation to preserve annual closure. This provides a controlled way to represent weather sensitivity while keeping demand exogenous.
- **Peak cap (optional):** If a peak cap D^{max} is configured, the generator enforces $D_{t,e}^{heat} \leq D^{max}$ and then re-scales to restore annual closure. If strict simultaneous enforcement of both a hard peak cap and exact annual closure is infeasible under the chosen logic, any residual is recorded explicitly (rather than silently absorbed) to preserve accounting transparency.

DemandPack is admitted to downstream evaluation only if all gates pass (tolerances are set in configuration and recorded in the run manifest):

1. **Timestamp integrity**

   UTC timestamps; strictly increasing; no duplicates; no missing hours; exactly $T = 8760$ for a standard (non-leap) year.

2. **Constant timestep**

   Δt inferred from timestamps equals 1 hour for all intervals.

3. **Non-negativity**

   $D_{t,e}^{heat} \geq 0 \; \forall \; t$, with explicit handling for numerical tolerances.

4. **Annual energy closure**

   Given target $E^{ann}[MW\mathrm{h}]$ (converted from GWh if needed). Enforce closure within tolerance ϵ_E:

$$\left| \sum_{t \in T} D_{t,e}^{heat} \Delta t - E^{ann} \right| \leq \epsilon_E$$

   If stochastic perturbations are enabled, the perturbed series is re-scaled to satisfy this identity.

5. **Optional peak cap (if enabled)**

   If a peak cap D^{max} is configured, enforce $D_{t,e}^{heat} \leq D^{max}$ and document how closure is preserved (or record any residual explicitly).

These gates ensure that any differences between pathways or scenarios arise from explicit scenario design (epoch tags, utility portfolios, uncertainty multipliers), rather than time misalignment, closure drift, or silent truncation. They operationalise the PoC's methodological claim that a coherent, auditable model can begin from limited information while remaining interface-stable.

In addition to the contractual CSV, DemandPack can generate diagnostic plots (annual time series, daily envelope, monthly totals, load duration curve, seasonal or hourly means, and distributions) to support transparent review of whether the synthetic series behaves consistently with the declared shaping assumptions. These diagnostics support interpretation and reporting, but do not alter the contractual artefact consumed by dispatch.

7.5.2 Epoch-tagged utility portfolios and availability representation

For each epoch tag e, the Edendale site is parameterised by an epoch-specific portfolio of utility units U_e (e.g., coal boilers, electrode boilers, biomass boilers). The purpose of epoch-tagging is to make structural change explicit and auditable, where differences between epochs arise from declared changes in the portfolio and parameters (commissioning, retirement, efficiency updates, carrier switching, or maintenance assumptions), rather than from implicit optimiser behaviour or hidden solver settings. Epoch tagging therefore functions as a structural provenance mechanism, ensuring that the "plant as-modelled" for any run can be reconstructed from version-controlled inputs and the *bundle/run* identifiers recorded in the output ledgers.

Unit parameterisation and carrier semantics

Each unit $u \in U_e$ is represented by a small set of transparent parameters sufficient for feasibility checks, carrier accounting, and cost/emissions ledgers:

- **Nameplate heat capacity** $K_{u,e}[MW_{th}]$
- **Thermal conversion efficiency** $\eta_{u,e}$ (dimensionless), interpreted consistently as "useful heat out per unit energy in"
- **Carrier label** $k(u) \in coal, biomass, electricity, \ldots$, used to route costing and emissions accounting

- **Cost parameters**, at minimum a carrier price $c_{u,e}^{carrier}$ on an explicit energy basis (e.g., NZD/MWh_{fuel}; NZD/GJ converted deterministically to MWh) with explicit unit conversions; and variable O&M if used) and optional variable O&M (NZD/MWh_{th})
- **Emissions parameters**, including an emission factor $\phi_u[tCO_2/MWh_{fuel}]$. Biomass treatment is declared explicitly (e.g., $\phi_u = 0$ for stack-neutral reporting; or a non-zero upstream proxy if adopted for sensitivity), so that cross-pathway comparisons are not driven by implicit conventions.

These parameters are provided via epoch-tagged utilities tables (utilities CSVs), so that the portfolio used in any run is fully determined by $(e, variant)$ and the *bundle/run* identifiers recorded in the run metadata. This is a deliberate "artefact-first" design to ensure the site-layer logic remain lightweight, while the input tables are the authoritative specification of the epoch portfolio.

Availability representation from maintenance windows

The PoC treats unit commitment, minimum up/down times, and ramping as out of scope for the reported results. Instead, operational realism is introduced through a deliberately simple, audit-friendly availability factor:

$$a_{u,t,e} \in [0,1] \quad ,$$

derived from epoch-tagged maintenance windows under *Input/site/maintenance*. For non-variant epochs (e.g., 2020, 2025, 2028), the pipeline resolves *maintenance_windows_<epoch>.csv*. For 2035 pathway epochs, it resolves *maintenance_windows_<epoch_tag>.csv* for example, *maintenance_windows_2035_EB.csv*, with a documented fallback to a generic 2035 file if present. If no maintenance file is found, all units default to full availability ($a_{u,t,e} = 1$); defaults are explicit and logged, rather than silently inferred.

Each maintenance record specifies (u, t^{start}, t^{end}, a), where $a = 0$ denotes outage and $a = 1$ denotes full availability (partial derates $0 < a < 1$ are admissible for future refinement). Windows are interpreted on the common hourly index, and *end_timestamp_utc* is treated as exclusive. The dispatch module converts the maintenance table into an hourly availability matrix aligned to the same UTC chronology as DemandPack and enforces it deterministically. If overlapping windows occur, the effective availability is taken as the minimum over applicable windows (conservative) and the overlap is logged as an input warning.

Effective available capacity and interpretability

The available heat capacity for unit u at hour t is:

$$\overline{K}_{u,t,e} = a_{u,t,e} K_{u,e}$$

This formulation is intentionally minimal, makes infeasibility attributable to a small set of inspectable causes (e.g., insufficient installed capacity, outage assumptions, or demand peaks), rather than solver tolerances. Because availability enters as a time-indexed bound on $\overline{K}_{u,t,e}$, it is invariant to the dispatch kernel choice (proportional, LP, or blockwise subset), preserving comparability across kernel swaps.

Importantly, availability does not alter the DemandPack contract. DemandPack remains a single heat-demand series; availability enters as separate, versioned tables that can be consumed by any dispatch kernel without expanding the thin-waist schema. This preserves modularity, where the same $a_{u,t,e}$ artefact can be used by proportional dispatch (reported PoC), the implemented LP kernel, or the implemented "optimal subset" kernel, without changing the interface between demand construction and dispatch evaluation.

Input integrity gates (site-side preconditions)

Before dispatch outputs are admitted to the thesis reporting bundle, the site inputs must pass basic integrity checks:

- **Portfolio completeness:** unique *unit_id* per epoch; required columns present; $K_{u,e} > 0$, $\eta_{u,e} > 0$ finite; carrier labels valid; cost and emissions fields either present or explicitly defaulted and logged.
- **Maintenance integrity:** windows are well-formed (UTC, valid start and end ordering); unit IDs match the utilities table; implied $a_{u,t,e} \in [0,1]$; overlaps are either prohibited or resolved by a declared rule; the generated hourly availability matrix is free of gaps/duplicates.
- **Chronology alignment:** availability matrix uses the common hourly UTC index as DemandPack, ensuring outages affect the correct hours and remain consistent across runs and pathway variants.

Together, these conventions establish an epoch-tagged, auditable mapping from "what assets exist and when they are available" to the capacity constraints that shape dispatch and, ultimately, the exported coupling signal consumed by the interface layer.

Unit capacities, efficiencies and emission factors for every epoch are listed in Appendix D (Table D.1). Annual heat demand is 819 GWh in all epochs, so differences between epochs arise only from the declared portfolio.

7.5.3 Proportional dispatch and ledger accounting (cost, emissions, feasibility)

Given hourly heat demand $D_{t,e}^{heat}$ from DemandPack and the effective available capacities $\overline{K}_{u,t,e}$, the PoC dispatch computes unit heat outputs $q_{u,t,e}[MW_{th}]$ and any unmet heat $q_{t,e}^{unserved}[MW_{th}]$ such that the hourly heat balance holds:

$$\sum_{u \in U_e} q_{u,t,e} + q_{t,e}^{unserved} = D_{t,e}^{heat} \quad , \qquad q_{t,e}^{unserved} \geq 0 \quad , \qquad 0 \leq q_{u,t,e} \leq \overline{K}_{u,t,e} \ .$$

The introduction of $q_{t,e}^{unserved}$ ensures feasibility under any portfolio and availability assumption, while making any shortfall explicit and auditable (rather than implicitly absorbed through truncation or hidden solver tolerances).

The site layer is a first-law model: it conserves energy in every hour. Each unit's fuel or electricity input is derived from the useful heat it delivers and its conversion efficiency, with all fuel energy expressed on a consistent lower-heating-value basis (Table 6.3). Direct emissions follow from the fuel each combustion unit uses. The heat balance closes in every hour, any shortfall is recorded explicitly as unserved heat, and annual totals reconcile with the DemandPack energy (Section 7.5.5). The model does not represent temperature levels, steam pressures or energy quality (exergy). This is a boundary choice, not an oversight. Temperature-resolved data were not available (Section 7.3.1), and both 2035 pathways deliver the same steam service, so a second-law analysis would add an energy-quality criterion without changing the cost and grid-interface comparison (Table 7.2). The DemandPack schema already reserves optional fields for temperature class and steam quality (Table 6.6), so this accounting can be added behind the same interface once such data become available.

Reported dispatch kernel and kernel swap-ability

Reported PoC results use a deterministic proportional allocation rule. After applying availability, demand is allocated across available units according to fixed shares (capacity-proportional and user-specified weights). Outputs are clipped at $\overline{K}_{u,t,e}$ with any residual reallocated to remaining units; any remaining deficit is recorded as $q_{t,e}^{unserved}$. This choice is deliberate and prioritises transparency, reproducibility, and causal attribution. Differences between epochs and pathway variants therefore arise from declared structural changes (portfolio composition, efficiencies, carrier assignments, maintenance assumptions), rather than from optimiser heuristics, solver settings, or non-observable tie-breaking behaviour.

While proportional dispatch is sufficient to substantiate the PoC's architectural claim (one-pass signalling with auditable artefacts), the site-layer implementation is intentionally structured to admit

optimisation-based dispatch and planning without changing the coupling surface. Concretely, dispatch is treated as a kernel-swappable evaluator operating behind an invariant artefact contract:

- Inputs remain unchanged: DemandPack (hourly demand), epoch-tagged utilities and availability, and exogenous SignalsPack fields used for costing and alignment.
- Outputs remain unchanged: hourly dispatch ledgers (long/wide), optional costed ledgers, annual KPI summaries, and the exported coupling signal $\Delta p_{t,e}^{elec}$.

This separation of interfaces (thin-waist artefacts) from internal solution method (dispatch kernel) is central to the framework's expansion capability. It allows the PoC to remain audit-ready under sparse data today, while remaining compatible with richer decision models (LP/MILP) when higher-fidelity inputs become available or when optimisation benefits outweigh added complexity.

Two swap-in optimisation kernels are implemented as capabilities under the same artefact schema, but are not required for the reported PoC bundle:

- Hourly LP dispatch kernel (economic allocation, no inter-temporal coupling), which solves each hour independently, allocating heat across units to minimise short-run variable cost (fuel and electricity and optional carbon proxy) while respecting capacity and heat balance. It introduces economically meaningful ordering (a merit-order analogue) without introducing behavioural assumptions about start-up, minimum up/down times, or ramping (at this stage). It is therefore a disciplined intermediate step between proportional allocation and full unit commitment.
- Blockwise "optimal subset" kernel (stylised commitment bridge towards MILP) introduces a coarse commitment representation by selecting, for each block (e.g., week), which subset of units is online, accounting for optional fixed online costs and start-up costs. Rather than solving a full-year MILP, the PoC implementation evaluates a finite candidate set of subsets per block and selects the lowest-cost feasible subset. This provides a forward-compatible pathway to full unit commitment while keeping the model inspectable and computationally manageable at PoC maturity.

These kernels are documented to demonstrate the framework's methodological headroom: optimisation can be integrated within the site layer (and later within regional electricity or biomass layers) without refactoring the artefact discipline or breaking coupling. Where optimisation becomes essential (e.g., when plant-level operational constraints, process integration, or supply-chain and logistics decisions materially affect cost and feasibility), the same DemandPack/SignalsPack contracts can be retained while the internal kernel is upgraded from proportional allocation to LP/MILP.

Carrier use, electricity exposure, and exported coupling signal

Each unit's energy input is computed from delivered heat and efficiency. For fuel-fired boilers:

$$f_{u,t,e} = \frac{q_{u,t,e}}{\eta_{u,e}} \left(MW_{fuel}\right) \ ,$$

and for electrode boilers (or other electrically driven heat supply):

$$p^{EB}_{u,t,e} = \frac{q_{u,t,e}}{\eta_{u,e}} \left(MW_e\right) \ .$$

These definitions keep the exported coupling signal well-defined and ensure electricity exposure is captured even when heat demand is the primary modelled quantity.

Ledger accounting: costs, emissions, and feasibility penalties

Costs are reported as additive components (fuel, electricity, carbon proxy, and penalties), with annual totals obtained by summing over t (with $\Delta t = 1$ hour). Denoting fuel price $c^{fuel}_{u,e}$ $[NZD/MWh_{fuel}]$ (or converted from NZD/GJ), electricity tariff $c^{elec}_{t,e}[NZD/MWh_e]$ supplied by SignalsPack (flat or time-varying), and an unserved penalty $c^{unserved}_e$ $[NZD/MWh_{th}]$,

$$C^{fuel}_e = \sum_t \sum_{u \in U^{fuel}_e} \left(f_{u,t,e} \Delta t\right) c^{fuel}_{u,e} \ ,$$

$$C^{elec}_e = \sum_t \left(\sum_{u \in U^{EB}_e} p^{EB}_{u,t,e} \right) \Delta t c^{elec}_{t,e} \ ,$$

$$C^{unserved}_e = \sum_t \left(q^{unserved}_{t,e} \Delta t\right) c^{unserved}_e \ .$$

The unserved penalty $c^{unserved}_e$ is a site-side feasibility diagnostic (heat not served). This is distinct from the regional interface VOLL used in the grid screening layer for electricity shedding. In the implementation, fuel prices may be specified as NZD/GJ (e.g., biomass) and are converted internally to a consistent energy basis before aggregation; the thesis reporting uses a single consistent basis to preserve comparability across carriers.

If an ETS-style carbon proxy is enabled, combustion emissions are computed from fuel use using emission factors ϕ_u (tCO_2/MWh_{fuel}), with carbon price $c^{CO_2}_e$ (NZD/tCO_2) applied consistently across epochs:

$$E^{CO_2}_e = \sum_t \sum_{u \in U^{fuel}_e} \left(f_{u,t,e} \Delta t\right) \phi_u \quad , \qquad C^{ETS}_e = c^{CO_2}_e E^{CO_2}_e \ .$$

The site-layer annual total is then:

$$C_e^{total} = C_e^{fuel} + C_e^{elec} + C_e^{ETS} + C_e^{unserved} \quad ,$$

with all components reported explicitly in the dispatch summary to avoid “black-box” aggregation.

A key reporting discipline in the PoC is that electricity accounting must remain comparable across pathways. Where electrode boiler electricity consumption appears as an implicit “fuel” in unit-level tables, the PoC additionally reports derived or effective electricity totals and costs (e.g., *annual_electricity_MWh_effective*, *annual_electricity_cost_nzd_effective*) so that EB–BB comparisons cannot be distorted by representation differences.

For each epoch/run, the dispatch module produces an auditable ledger set: an hourly long-form ledger (unit-by-hour heat output and carrier use), an hourly wide-form ledger (system totals and per-unit outputs), an optional costed long-form ledger (cost attribution), and an annual summary table used as the KPI source for Chapter 8. Because these outputs are invariant to kernel choice, future upgrades (LP/MILP kernels, richer unit constraints, additional carriers, or supply-chain terms) can be introduced without changing the thin-waist coupling artefacts or the reporting pipeline.

7.5.4 Exported coupling signal

The site-to-region coupling variable in the PoC is the incremental electricity demand attributable to electrified heat supply (EB operation), exported as an hourly series in MW on the same timestamps as DemandPack and SignalsPack. Operationally, for epoch e,

$$\Delta p_{t,e}^{elec} = \sum_{u \in U_e^{EB}} p_{u,t,e}^{EB} \ .$$

This signal is exported as *incremental_electricity_MW_<epoch_tag>.csv* and is interpreted as the additional electricity demand at the site boundary attributable to electrified heat production, under the site-layer accounting conventions described above.

The site module intentionally does not net this signal against a full site electricity baseline. Instead, the interface layer combines an (exogenous) baseline GXP loading representation with $\Delta p_{t,e}^{elec}$ for hosting-capacity screening and upgrade exposure. This design keeps the coupling contract minimal and stable, and it supports extension to multiple sites or alternative baseline definitions at the regional layer without modifying the site dispatch model.

7.5.5 Site-layer validation checks used as reporting gates

Before an epoch run is admitted to the thesis reporting bundle, the site layer must satisfy a small set of deterministic validation gates applied to the canonical artefacts produced by the PoC pipeline (DemandPack manifest and hourly demand series; dispatch hourly ledgers; dispatch annual summary; and, where relevant, the exported incremental-electricity signal). These gates implement the "acceptance checklist" philosophy embedded in the repository smoke tests and ensure that all reported KPIs are reproducible from inspectable, time-aligned ledgers.

Gate 1 — Chronology and schema integrity (time consistency).

All hourly artefacts must conform to the canonical 8760-hour UTC chronology and valid schemas: required columns exist; values are finite (no NaN/inf); and timestamps contain no gaps or duplicates. In addition, any derived matrices used by dispatch (e.g., availability multipliers constructed from maintenance windows) must align exactly to the same timestamp index. These checks prevent subtle errors where misaligned time bases can create spurious closure, cost, or exceedance behaviour in later coupling steps.

Gate 2 — Heat-balance feasibility and energy closure (physical consistency).

For every hour t, the dispatch ledger must satisfy the heat-balance identity:

$$\sum_{u \in U_e} q_{u,t,e} + q_{t,e}^{unserved} = D_{t,e}^{heat} \quad ,$$

with unit outputs bounded by effective available capacities and non-negativity enforced. Annual totals must reconcile to within a declared tolerance between (i) the DemandPack annual target and (ii) the integral of the hourly demand series, and between (iii) hourly dispatch totals and (iv) annual summary totals. Any unserved heat must be reported explicitly (via both hourly ledgers and an annual UNSERVED row) and must never be silently absorbed through scaling or residual adjustments.

Gate 3 — Accounting reconciliation (ledger completeness and additive cost identity).

Annual KPIs reported in the dispatch summary must reconcile exactly to the underlying hourly ledgers: fuel and electricity inputs are consistent with delivered heat and efficiencies; carrier totals are internally consistent; and cost components sum to the reported annual total (e.g., $C^{fuel} + C^{elec} + C^{carbon} + C^{unserved} = C^{total}$). The summary file must contain the expected aggregation rows (at minimum TOTAL, and UNSERVED where applicable) to ensure that downstream reporting and comparison scripts operate on a stable schema.

For the “thesis-ready” reporting bundle, the PoC applies a strict release rule consistent with the repository smoke tests; required artefacts exist, TOTAL/UNSERVED aggregations are present, and unserved energy and penalty costs are zero unless a run is explicitly designated as a stress-test case and labelled as such in the results narrative. This rule preserves comparability across epochs and pathway variants and avoids accidental inclusion of infeasible runs as if they were candidate strategies.

Table 7.4 lists the exact artefacts to which the site-layer reporting gates are applied, using the bundle epoch–run identifier to preserve provenance and prevent cross-run contamination.

Table 7-4 Site-layer reporting gates and file contracts

Gate	Artefact	Path pattern	Checks (minimum)
G1	DemandPack manifest	*…/epoch<e>/demandpack/**/meta/run_manifest.json*	Present; provenance recorded; closure metadata populated.
G1	Hourly heat demand	*…/epoch<e>/demandpack/demandpack/hourly_heat_demand_<e>.csv*	8760 UTC; no gaps/dupes; finite; ≥0; schema valid.
G2	Dispatch ledger (long)	*…/epoch<e>/dispatch/<runid>/site_dispatch_<e>_long.csv*	Feasible ($0 \le q \le \overline{K}$); timestamps aligned; UNSERVED explicit.
G2	Dispatch ledger (wide)	*…/epoch<e>/dispatch/<runid>/site_dispatch_<e>_wide.csv*	Aggregations consistent with long; no missing hours; finite.
G3	Costed ledger	*…/epoch<e>/dispatch/<runid>/site_dispatch_<e>_long_costed.csv*	Inputs consistent with (η); cost fields finite; roll-up matches summary.
G3	Dispatch KPIs (summary)	*…/epoch<e>/dispatch/<runid>/site_dispatch_<e>_summary.csv*	Required fields; TOTAL/UNSERVED rows; totals reconcile; components sum to annual_total_cost_nzd.
G4*	Coupling signal (ΔE)	*…/epoch<e>/dispatch/<runid>/signals/incremental_electricity_MW_<e>.csv*	8760 UTC; aligned; sign convention; ≥0; plausibility bounds.
G0	Bundle smoke test	*…/runs/<bundle>/*	Completeness across epochs; schemas stable; no NaNs/inf; release rule for unserved/penalties; 2035 EB/BB both present for paired comparison.

*G4 applies only to electrified epochs (e.g., 2035_EB; and to 2035_BB only to the extent electrified units exist in the committed prefix).

Together, these gates are not merely technical hygiene; they are the methodological backbone of the PoC claim. They guarantee that subsequent regional screening and robustness calculations operate on signals and KPIs that are time-consistent, auditable, and comparable across pathway variants, and that any infeasibility is an explicit, inspectable outcome rather than an artefact of data misalignment or hidden aggregation.

7.6 Exploratory layer: screening and robustness evaluation

This section defines the PoC's exploratory layer for testing the robustness of the site's 2035 pathway choice under uncertainty. Following the committed mid-2020s electrification steps, the PoC frames a single strategic fork for 2035: replacing the remaining coal-derived heat supply either through further electrification or via a biomass pathway, yielding two directly comparable variants (2035_EB and 2035_BB). Given limited defensible high-frequency site operational data, and several material dynamics that cannot yet be parameterised credibly (e.g., unit commitment, production-driven variability, tariff response, and process-resolved constraints), the PoC retains an auditable deterministic site core and uses the exploratory layer to stress the most binding uncertainties and to prioritise data improvements with the highest decision value.

The dominant feasibility driver lies largely outside the site boundary and is governed by grid hosting capacity, represented by hourly headroom at the Edendale GXP, together with the potential requirement for discrete network upgrades. The Southland RETA assessments indicate limited spare capacity at this interface and provide screening-level upgrade options suitable for comparative analysis [44]. The resulting feasibility constraint exhibits threshold behaviour, where modest reductions in headroom or increases in incremental load can create exceedance hours that trigger either load shedding (valued using VOLL) or selection of a higher-capacity upgrade option. Because this binding constraint is located in the grid layer, the PoC applies RDM-style screening within the regional proxy layer, which evaluates a discrete upgrade menu across many plausible futures using the site's exported incremental electricity trace under a paired set of uncertainty realisations.

The site model remains a validated deterministic core (DemandPack → proportional dispatch → annual summaries), while deep uncertainty is explored in the grid screening layer, where feasibility is binding at hourly resolution. To extend robustness assessment back to the site pathway decision without breaking one-pass coupling, a post-processing robustness overlay is applied on the same paired futures set. This overlay perturbs key site-side cost components (electricity price, biomass price/scarcity proxy, and a carbon-price proxy) and computes decision-relevant indicators (e.g., regret and satisficing rates for 2035_EB versus 2035_BB) without re-solving dispatch.

This separation is an explicit PoC scope choice rather than a limitation of the modular architecture. A full site-level RDM analysis would re-run (and potentially re-optimise) the site dispatch for each future. In the present PoC, where dispatch is proportional and exhibits limited behavioural adaptation, re-solving would add limited additional insight for purely multiplicative cost uncertainties, except where uncertainties materially change feasibility (e.g., demand growth pushing above capacity, outages/deratings, binding fuel availability constraints). Full site-RDM becomes materially meaningful

once the site layer adopts a cost-aware dispatch kernel (e.g., LP/MILP) and uncertainties act through binding operational constraints (tariff shapes, deratings, biomass caps, outage schedules), and/or once additional decision levers are introduced (sizing, storage, hybridisation, demand response). The PoC therefore implements core elements of a DMDU workflow, decision definition, ensemble exploration, performance evaluation, and robustness metrics, while deferring vulnerability diagnosis (scenario discovery) and adaptive policy design (DAPP and MORDM) to future work enabled by the same thin-waist interfaces.

7.6.1 Decision locus and decision objects

The PoC is structured around a clear separation between where the strategic decision is taken and where binding feasibility is evaluated. The decision locus is the industrial site, where the decision concerns the long-run transition pathways for process heat supply in 2035. Both pathways share the committed pre-2035 investments and differ only in the post-commitment treatment of the remaining coal boiler capacity. The site model maps this pathway choice to a coupling signal, an hourly incremental electricity series attributable to electrified heat supply, aligned to the common UTC hourly index used throughout the PoC.

Formally, the site-level decision is represented as a discrete pathway choice (implemented as *epoch_tag*):

$$p^{site} \in \{\text{EB,BB}\}$$

In contrast, the regional layer is not treated as a full planning model. It is a screening proxy that tests whether the incremental electricity trace can be hosted within an assumed headroom envelope and, if not, what discrete reinforcement option would be indicated under the screening rule. The regional decision object is therefore defined over a finite upgrade menu (implemented in *rdm_matrix_**):

$$p^{grid} \in \{s_0, s_1, \dots, s_K, s_{\text{AUTO}}\}$$

where s_0 denotes a no-upgrade baseline and $s_1, \dots, s_K$ denote discrete upgrade options parameterised by indicative capacity increments and costs consistent with screening-level planning evidence. The option s_{AUTO} is not treated as a real strategy in p^{grid}, instead, it is used as a benchmark selector and is interpreted purely as an internal reference for regret-style comparisons rather than as an operational or institutional decision rule.

7.6.2 Thin-waist artefacts and regional signals (screening inputs)

The exploratory layer is enabled by a small set of thin-waist artefacts that carry information from the deterministic site core to the regional proxy (and into the robustness overlay) without exposing internal model detail. These artefacts define the screening inputs used to test grid-hosting feasibility and to evaluate discrete upgrade strategies under uncertainty, while avoiding exposure of internal site model detail. Table 7.5 summarises the artefacts, their minimal schemas, and their roles in the screening workflow.

Table 7-5 Artefact table (screening inputs)

Artefact	Key fields	Producer → consumer	Screening role
ΔE_t incremental electricity	*timestamp_utc, incremental_electricity_MW*	site_dispatch → *run_rdm_2035*	Coupling signal; incremental load to be screened against headroom.
H_t headroom at GXP	*timestamp_utc, headroom_MW*	CSV or generator → *run_rdm_2035*	Hosting envelope; defines binding hours and exceedance.
Upgrade menu (strategy set)	*name, capacity_MW, capex_nzd*	TOML → *run_rdm_2035*	Discrete reinforcement options (incl. "none"); supports AUTO benchmark.
Tariff signal (site costing input)	*Flat/timestamp_utc, elec_price_nzd_per_MWh*	signals config/CSV → *site_dispatch*	Site costing only; not updated by screening (no congestion feedback).

The principal coupling signal is the site-exported incremental electricity trace ΔE_t which represents electrification-driven incremental grid demand on the common hourly UTC index. Regional feasibility is assessed by comparing ΔE_t against an hourly headroom envelope H_t at the Edendale GXP. Where ΔE_t exceeds H_t, the regional proxy records exceedance hours and implied shedding (valued using VOLL within the screening layer), and evaluates whether a discrete reinforcement option from the upgrade menu would be preferred under the screening rule.

Headroom and tariffs are treated as exogenous signals within the PoC scope boundary. If a measured (or externally generated) headroom time series is available, it is used directly; otherwise, the pipeline uses a deterministic PoC headroom generator and issues an explicit warning to indicate that results are conditional on stylised assumptions. Electricity tariffs used in the site summaries are supplied as either a flat value (from configuration) or a time-varying ToU series but are not updated by regional screening results. This prevents circular dependence and preserves the one-pass coupling claim.

Upgrade options are treated as screening-level discrete alternatives with indicative capacity and capex suitable for comparative robustness evaluation. Any recorded attributes such as lead time are treated as metadata unless explicitly represented in the screening equations (i.e., the PoC does not simulate

construction timing or phased delivery unless stated). These choices imply a clear interpretive guardrail: screening outputs are conditional on the assumed headroom proxy and upgrade menu parameters and should not be read as detailed network planning or market outcomes.

All artefacts used in screening are aligned at hourly resolution on the same UTC timestamp index, consistent with DemandPack and dispatch outputs. This ensures that exceedance hours, shed calculations, and annualised screening metrics are computed from a consistent temporal basis and remain directly comparable across pathways under the paired futures contract.

7.6.3 Uncertainty design and paired-futures contract

This subsection specifies the uncertainty set and the pairing contract required to support fair, like-for-like comparison between the two site pathway strategies under deep uncertainty. Consistent with Chapter 5, uncertainty is represented by a set of plausible futures Ω, without asserting probabilities as decision-relevant forecasts. Each future $\omega \in \Omega$ is an internally consistent realisation of exogenous drivers that affect grid hosting feasibility and the costs of remedial actions within the screening proxy.

Let $\Omega = \{\omega_0, \omega_1, \dots, \omega_{N-1}\}$ denote the ensemble of N futures used in the grid screening experiments. Futures are indexed by a unique identifier *future_id*, such that ω_f denotes the future associated with *future_id = f*. The PoC treats *future_id* as the join key across all exploratory outputs, ensuring that results from different layers can be compared and aggregated without ambiguity.

For each future ω, the grid screening layer uses a vector of uncertainty parameters $U(\omega)$ defined in the futures file:

$$U(\omega) := \ (U_{head}(\omega),\ U_{\Delta E}(\omega),\ U_{capex}(\omega),\ U_{VOLL}(\omega),\ U_{cons}(\omega)) \ ,$$

where:

- $U_{head}(\omega)$ (headroom multiplier) captures uncertainty in the available GXP hosting envelope;
- $U_{\Delta E}(\omega)$ (incremental-load multiplier) captures uncertainty in electrification intensity and/or incremental demand growth attributable to the site pathway;
- $U_{capex}(\omega)$ (capex multiplier) captures uncertainty in upgrade capital costs at screening level;
- $U_{VOLL}(\omega)$ (value of lost load) captures uncertainty in the valuation of unserved energy used in screening;
- $U_{cons}(\omega)$ (consents uplift) captures uncertainty in consenting or delivery overheads applied to upgrade costs.

These uncertainty factors are applied consistently across all evaluated strategies and are treated as exogenous drivers within the PoC scope boundary. Their operational use (e.g., how they modify signals and cost components) is defined in the screening mechanics in Section 7.6.4.

The paired-futures contract ensures that differences between the pathway outcomes are attributable to the pathway choice rather than to different uncertainty realisations. Formally, the contract requires that both site strategies $p^{site} \in \{\text{EB,BB}\}$ are evaluated over the same future set Ω. In practice this means:

1. The grid screening outputs for 2035_EB and 2035_BB must contain an identical set of *future_id* values.
2. Any comparison table constructed across pathways (by *future_id*) is restricted to this shared set, ensuring one-to-one pairing ω_f across pathways.
3. The futures definition used for a reporting bundle is treated as canonical for that bundle, so that all downstream comparisons and robustness summaries refer to a single, auditable uncertainty envelope.

This pairing requirement operationalises the Chapter-5 RDM principle that performance comparisons $J(p, \omega)$ must be evaluated on a common uncertainty basis when computing regret or satisficing rates. In the PoC, the pairing contract is enforced via explicit validation checks that fail if the *future_id* sets differ between pathway runs, preventing non-paired comparisons from entering thesis reporting bundles.

Baseline prices, emission factors, the upgrade menu and the screening assumptions to which the uncertainty multipliers are applied are listed in Appendix D (Table D.2); the ranges explored across the 100 futures are reported in Table 8.8.

7.6.4 Regional screening mechanics and robustness metrics

This subsection operationalises the regional proxy layer introduced in Sections 7.6.2–7.6.3 by defining the screening mechanics used to evaluate (i) grid-hosting feasibility at the Edendale GXP and (ii) the relative attractiveness of discrete reinforcement options under the futures ensemble Ω. The screening is intentionally positioned as an interface-level robustness screen rather than a planning model: it consumes a thin-waist coupling signal (the pathway-specific incremental electricity trace) and evaluates hosting feasibility against an exogenous headroom envelope and a finite upgrade menu, without any feedback to site dispatch.

The screening takes as inputs the pathway-specific incremental electricity trace ΔE_t(MW), the GXP headroom envelope H_t(MW), and an upgrade menu $\mathcal{P}^{grid} = \{S_0, \dots, S_K\}$. These artefacts are aligned on the common hourly UTC index (*timestamp_utc*) to ensure binding-hour identification and annual aggregation are computed on a consistent temporal basis. In the PoC, headroom and tariffs are treated as exogenous signals. When measured or externally generated headroom is unavailable, a deterministic headroom generator is used with an explicit warning that results are conditional on stylised assumptions.

The upgrade menu is parameterised using RETA's Southland network findings, which identifies Edendale as a "major complexity" connection and provides indicative capacity increments and capital costs suitable for screening-level comparison (e.g., +21 MW, +32 MW, +97 MW), with a further high-capacity option added in the PoC as an explicitly flagged assumption to span the electrification envelope.

A deliberate corollary of this interface-level framing is that the screening is treated as an epoch-scoped assessment of hosting exposure for the pathway's exported trace, rather than as a sequential network investment plan that carries forward reinforcement decisions from earlier epochs. In other words, the upgrade menu $\mathcal{P}^{grid}$ is evaluated against the assumed headroom envelope H_t for the screened epoch, and the resulting upgrade costs should be interpreted as the screening-level reinforcement requirement implied by the 2035 pathway under the assumed proxy signals, not as "additional" reinforcement conditional on an explicitly modelled 2028 upgrade state. This scope choice is methodologically valuable and preserves one-pass coupling, avoids introducing recourse and state variables into the PoC, and keeps the thin-waist artefact contract (ΔE_t, H_t, $\mathcal{P}^{grid}$) transparent and reproducible. Sequential consistency is not precluded by the architecture; future staged formulations (multi-epoch recourse) can be introduced by modifying the exogenous headroom signal (or pre-applying an initial ΔK) to represent prior reinforcements, without changing the screening equations or the site-to-region interface logic.

Each future $\omega \in \Omega$ applies the uncertainty multipliers $U(\omega)$ defined in Section 7.6.3 (e.g., headroom, incremental-load, capex, VOLL, consents uplift). These parameters are treated as exogenous scenario descriptors rather than forecasts and are applied consistently across all strategies for each ω.

Signal transformation under uncertainty

For a given future ω, the effective incremental demand and headroom are:

$$\Delta E_t(\omega) \;=\; U_{\Delta E}(\omega)\,\Delta E_t, H_t(\omega) \;=\; U_{\mathrm{head}}(\omega)\,H_t \quad ,$$

where $\Delta E_t$and $H_t$are the baseline time series at hourly resolution (UTC).

Each upgrade strategy $p^{grid} \in \mathcal{P}^{grid}$ provides an incremental hosting capacity $\Delta K(p^{grid})$(MW), with $\Delta K(S_0) = 0$. The strategy-specific hosting envelope is therefore:

$$\bar{H}_t(\omega, p^{grid}) \;=\; H_t(\omega) \;+\; \Delta K(p^{grid}) \;.$$

In the PoC pipeline, the incremental electricity trace is exported by the site layer and passed to *src.run_rdm_2035* (via the run *bundle/run-id* conventions), while headroom is supplied either from an input CSV or from the deterministic headroom generator when required.

Exceedance and shed calculations

Hourly exceedance (MW) is computed as:

$$x_t(\omega, p^{grid}) \;=\; \max\big(0, \Delta E_t(\omega) - \bar{H}_t(\omega, p^{grid})\big) \;.$$

With $\Delta t = 1$hour, annual shed energy (MWh) and peak exceedance (MW) are:

$$Shed(\omega, p^{grid}) \;=\; \sum_t x_t\,(\omega, p^{grid})\,\Delta t, \qquad x^{\max}(\omega, p^{grid}) \;=\; \max_t x_t(\omega, p^{grid}) \;.$$

The number of binding hours is recorded as

$$n_{\text{bind}}(\omega, p^{grid}) = \textstyle\sum_t \mathbf{1}\{\, x_t(\omega, p^{grid}) > 0\} \;.$$

Where required for reporting, the shed fraction is defined as:

$$\phi(\omega, p^{grid}) \;=\; \frac{Shed(\omega, p^{grid})}{\sum_t \Delta\, E_t(\omega)\,\Delta t} \;.$$

Here, x_t represents interface-level unserved incremental demand implied by the headroom constraint: it is not a detailed operational security simulation (e.g., contingency analysis), and it does not distinguish between curtailment, dispatchable demand response, or network operational actions. The key purpose is to identify (a) whether hosting becomes binding, (b) how frequently, and (c) how severe the implied exceedance is (energy and peak). The PoC records these exceedance-derived quantities directly in the RDM outputs, including *annual_shed_MWh*, *shed_fraction*, *unserved_peak_MW*, and *n_hours_binding*, computed by evaluating each (future, strategy) combination across the hourly traces.

Upgrade cost annualisation and total screening cost

Let $Capex(p^{grid})$ denote the nominal capex for strategy p^{grid}(NZD), with $\text{Capex}(S_0) = 0$. Screening-level uncertainty is applied multiplicatively:

$$\widetilde{\text{Capex}}(\omega, p^{grid}) \;=\; \text{Capex}(p^{grid}) \cdot U_{\text{capex}}(\omega) \cdot U_{\text{cons}}(\omega) \;.$$

The annualised upgrade cost is calculated using a capital recovery factor $\mathrm{CRF}(r, n)$with discount rate $r$and asset life n:

$$AUC(\omega, p^{grid}) \;=\; CRF(r,n) \cdot \widetilde{\mathrm{Capex}}(\omega, p^{grid}), CRF(r,n) = \frac{r}{1-(1+r)^{-n}} \;.$$

Shed cost uses the future-specific value of lost load $U_{\mathrm{VOLL}}(\omega)$(NZD/MWh):

$$SC(\omega, p^{grid}) \;=\; U_{\mathrm{VOLL}}(\omega) \cdot Shed(\omega, p^{grid}) \;.$$

The screening objective (total annual cost) is then:

$$J^{grid}\big(p^{grid}, \omega\big) \;=\; AUC(\omega, p^{grid}) \;+\; SC(\omega, p^{grid}) \;.$$

This objective deliberately collapses regional feasibility exposure into two transparent components: (i) annualised reinforcement expenditure (with consenting/delivery uplift treated as a multiplicative uncertainty), and (ii) an economic proxy for residual infeasibility valued using VOLL. This is aligned with the PoC's stated boundary: a tractable decision screen at the interface, not a representation of price formation or network investment optimisation. The uncertainty multipliers used in the futures file explicitly include *U_upgrade_capex_mult*, *U_consents_uplift*, and *U_voll*, and the outputs report annualised upgrade cost, annual shed cost, and their sum (*total_cost_nzd*).

Satisficing constraints (screening feasibility)

The screening layer labels a strategy as satisficing in a future if it meets configured feasibility tolerances. In the PoC this is implemented as a Boolean pass/fail flag consistent with constraints of the form:

$$\mathbf{1}\{Shed(\omega, p^{grid}) \le \bar{Shed} \;\wedge\; x^{\max}(\omega, p^{grid}) \le \bar{x} \;\wedge\; \phi(\omega, p^{grid}) \le \bar{\phi}\} \;,$$

where $\bar{Shed}, \bar{x}, \bar{\phi}$ are screening thresholds (with $\bar{Shed} = 0$ representing "no shedding allowed", up to numerical tolerance). he PoC implements satisficing with explicit tolerances (e.g., *shed_MWh_max*, *max_exceedance_MW*, and optionally *shed_fraction_max*) to avoid misclassifying numerical noise as infeasibility.

Satisficing in this layer should be read as "passes the configured screening guardrails given the assumed headroom proxy and upgrade menu" rather than a formal reliability or security standard. This is important for thesis defensibility: the screen is designed to flag exposure and compare discrete options, not to certify compliance with network planning standards.

Benchmark selection and regret (RDM-style comparison)

To mirror the regret framing in Chapter 5, the screening uses a per-future benchmark selector S_{AUTO} that chooses the least-cost upgrade within the menu:

$$p_{AUTO}^{grid}(\omega) \;=\; \arg\min_{p^{grid}\in\mathcal{P}^{grid}} J^{grid}\left(p^{grid},\omega\right).$$

Regret relative to this benchmark is computed as:

$$r^{grid}\left(p^{grid},\omega\right) \;=\; J^{grid}\left(p^{grid},\omega\right) \;-\; J^{grid}\left(p_{AUTO}^{grid}(\omega),\omega\right).$$

S_{AUTO} is an internal reference device that makes trade-offs legible across futures, and answers, "what would the screening rule pick in this future if it were allowed to choose freely from the menu?", enabling regret distributions to be plotted and compared. It is not interpreted as an institutional rule or as an operational dispatch decision.

Outputs

Two artefact families produced by src.run_rdm_2035 support the Chapter 8 results narrative:

1. **RDM summary** (one row per ω, benchmark-selected upgrade):

   Path templates include *Output/runs/<bundle>/rdm/rdm_summary_2035_EB.csv* and *..._2035_BB.csv*, generated by running the RDM screen separately for each pathway (with identical futures). This file supports:

   - Upgrade selection frequency (how often each option is selected by
   - Distributions of annual_shed_MWh, n_hours_binding, unserved_peak_MW, and total_cost_nzd;
   - Pathway-to-pathway comparisons under paired futures (EB vs BB, by future_id).

2. **RDM matrix** (all p^{grid} across all ω):

   The long-form matrix rdm_matrix_<epoch_tag>.csv supports:

   - Regret distributions $r^{grid}(p^{grid},\omega)$ and pass rates (satisficing frequency) by strategy;
   - Trade-offs between upgrade cost and shed cost across the entire menu (including identification of dominated strategies under the configured tolerances).

Finally, a paired-futures comparison CSV (*rdm_compare_2035_EB_vs_BB.csv*) is generated after both pathway screens, providing a direct one-to-one join by *future_id* for the principal EB–BB comparative plots in Chapter 8.

7.6.5 Site decision robustness overlay

This subsection defines a post-processing robustness overlay that extends the uncertainty evaluation from the regional screening decision to the site pathway decision $p_{\text{site}} \in \{EB, BB\}$. Concretely, the overlay does not re-solve the site dispatch and does not introduce feedback from the regional proxy layer back into the site core; it operates solely by re-costing a small set of annual cost components using per-future multipliers, then adding the pathway-consistent regional screening cost. This is consistent with the PoC methodological scope and the deliberate separation of concerns between the site and regional layers.

Inputs (per bundle, paired futures)

For each pathway p_{site}, the overlay consumes three artefact families produced earlier in the pipeline:

1. **Dispatch TOTAL annual summary (site core):** The overlay reads the *TOTAL* row from the dispatch summary for each pathway epoch-tag (e.g., 2035_EB, 2035_BB). This provides the baseline annual cost components and emissions that the overlay perturbs, without revisiting hourly dispatch quantities.

2. **Regional screening summary (grid proxy):** From the RDM screening output, the overlay uses the benchmark-selected grid cost $J^{grid}(p^{site}, \omega)$ evaluated on the incremental electricity trace ΔE_t exported by that pathway. In implementation terms, this corresponds to the per-future "total cost" reported for the benchmark selection, i.e. upgrade + shed screening cost for each *future_id*.

3. **Canonical futures table (paired):** The overlay uses a bundle-scoped canonical futures table containing both the grid-side multipliers $U(\omega)$ and the site-side multipliers $M(\omega)$. The canonical table is the mechanism that guarantees paired-futures comparability across EB and BB.

The site-side multipliers are:

$$M(\omega) = (P_{elec}(\omega), P_{bio}(\omega), P_{ETS}(\omega)) \ ,$$

representing electricity price, biomass delivered cost (scarcity proxy), and an ETS (carbon-price proxy), respectively. In the PoC artefacts these correspond directly to the canonical futures columns *P_elec_mult*, *P_biomass_mult*, and *ETS_mult*.

Perturbation model (component-wise future costing)

Let the baseline site annual cost for pathway p_{site} be decomposed as:

$$J_0^{site}(p^{site}) \ = \ C_0^{elec}(p^{site}) \ + \ C_0^{bio}(p^{site}) \ + \ C_0^{ETS}(p^{site}) \ + \ C_0^{other}(p^{site}) \ ,$$

where C_0^{other} aggregates remaining components held fixed in the overlay (e.g., any non-biomass fuel costs and other operating-cost terms as represented in the PoC annual summaries).

The future-costed site component is:

$$J^{site}(p^{site}, \omega) = P_{\text{elec}}(\omega)\, C_0^{elec}(p^{site}) + P_{\text{bio}}(\omega)\, C_0^{\text{bio}}(p^{site}) + P_{\text{ETS}}(\omega)\, C_0^{\text{ETS}}(p^{site})\, C_0^{\text{other}}(p^{site}) .$$

Operationally, the PoC applies these multipliers post hoc to the corresponding annual cost fields from the dispatch *TOTAL* summaries (electricity, biomass fuel, carbon), leaving all other reported components unchanged. This aligns with the explicit "reporting overlay" boundary: the multipliers alter *accounting* outcomes, not dispatch feasibility or hourly allocations.

The overlay then forms a combined (system) annual performance measure by adding the pathway-consistent regional screening cost:

$$J^{sys}(p^{site}, \omega) \;=\; J^{site}(p^{site}, \omega) \;+\; J^{grid}_{AUTO}(p^{site}, \omega) \;,$$

where $J^{grid}_{AUTO}(p^{site}, \omega)$ is the benchmark-selected screening cost from Section 7.6.4, evaluated using the incremental electricity trace produced by that pathway (and therefore permitted to differ between EB and BB even under the same ω).

Robustness metrics for the site pathway decision

Using the Chapter 5 yardsticks, the overlay reports a deliberately lean set of decision-facing indicators for $p_{\text{site}} \in \{EB, BB\}$, computed over the paired futures Ω:

- Win rate (paired-futures dominance):

$$W(\text{EB}) = \frac{1}{\mid \Omega \mid} \sum_{\omega \in \Omega} \mathbf{1}\{J^{sys}(\text{EB}, \omega) \leq J^{sys}(\text{BB}, \omega)\}, W(\text{BB}) = 1 - W(\text{EB}) \;.$$

- Two-alternative regret (paired, non-negative by construction):

$$r^{site}(p^{site}, \omega) = J^{sys}(p^{site}, \omega) - \min_{p' \in \{\text{EB},\text{BB}\}} J^{sys}(p', \omega) \;.$$

Regret is summarised using distributional statistics (e.g., median, P90) and visualised via a regret CDF.

Satisficing rate for a configured threshold θ (e.g., a system-cost ceiling; optionally complemented by emissions or reliability-style thresholds if enabled):

$$S_\theta(p^{site}) = \frac{1}{\mid \Omega \mid} \sum_{\omega \in \Omega} \mathbf{1}\{J^{sys}(p^{site}, \omega) \leq \theta\} \;.$$

These metrics are produced as part of the PoC robustness consolidation step, which joins (i) pathway dispatch summaries, (ii) pathway RDM screening outputs, and (iii) the canonical paired futures table, then computes J^{sys} and the above indicators without modifying upstream artefacts. For thesis reporting, two canonical figures are admitted: (i) a regret CDF comparing EB and BB, and (ii) a system-cost distribution plot (e.g., boxplot) over paired futures.

Interpretation and approximation boundary

Under proportional dispatch (used for reported PoC results), this overlay is a close approximation to a full re-run for uncertainties that act *multiplicatively* on annual costs without changing feasible operations (electricity price scaling, biomass price scaling, and ETS price scaling). In such cases, the overlay preserves the PoC's methodological intent and expresses *decision sensitivity* to deep uncertainty while retaining one-pass coupling and auditability.

Divergence is expected when uncertainties would change feasibility or operational choices, effects not represented by proportional dispatch or by post hoc costing, such as demand growth pushing above available heat capacity; unit deratings (outages) that become binding; explicit biomass availability constraints; time-varying tariff shapes that induce behavioural shifts; or additional site decision levers (storage, hybridisation, demand response). These conditions define the methodological boundary at which a full site-level RDM (i.e., re-solving dispatch per ω with a cost-aware kernel) becomes necessary to capture decision-relevant dynamics, rather than simply re-pricing an invariant operating pattern.

7.7 Integration and reporting gates

This section defines the integration gates that must be satisfied before any Proof-of-Concept (PoC) outputs (tables and figures) are admitted to the reporting bundle. The intent is to ensure that results are (i) traceable to canonical artefacts, (ii) time-consistent across the site–interface–screening chain, and (iii) comparable across pathway variants under a paired uncertainty set. The underlying validation checks and artefact contracts are specified earlier (notably the site-layer validation gates and the minimum artefact register); here they are consolidated into a release rule for Chapter 8 reporting.

All figures and tables presented in Chapter 8 draw only from a bundle-tagged run directory in which the canonical artefacts listed in the minimum register exist, resolve locally (i.e., without machine-specific scratch paths), and pass their acceptance gates. The compact register in Table 7.1 provides the admissibility criterion, and if any required artefact is missing or fails its recorded checks, the corresponding run is excluded from the reporting set, rather than repaired ex post in the analysis layer.

Two additional governance requirements apply at the interface-to-screening boundary and to cross-pathway comparisons. First, paired-futures fairness is enforced, so any EB–BB comparison must be constructed by an explicit join on *future_id*, using a common Ω (plausible futures). If either pathway is missing any future ensembles ω, the comparison is either restricted to the paired subset (with exclusions recorded) or rejected as not thesis-ready; unpaired distributional comparisons are not permitted. Second, closed interface accounting is required for both screening and overlay results: grid screening outputs and the reporting-only site robustness overlay are admitted only when their internal closure identities (Sections 7.6.4–7.6.5) hold within tolerance, preventing robustness metrics from being computed on inconsistent cost components.

Operationally, the release rule is implemented through (i) per-artefact schema (chronology) checks, (ii) bundle-level validation of paired futures and required columns, and (iii) a curated reporting pack (*thesis_pack*/…) containing read-only derivatives and a pointer map to canonical source artefacts. With these governance conditions satisfied, the thesis can now turn from *how evidence is generated and admitted* to *what that evidence shows*.

Chapter 8
Results and Discussions

8.1 Introduction

This chapter presents the results and discussion for the Edendale–Southland proof-of-concept (PoC). Its purpose is to demonstrate the framework mechanics as an auditable starting point for an iterative exploratory testbed: a baseline implementation that can be progressively strengthened (e.g., richer data, expanded futures, and additional uncertainties) to increase decision confidence over time. Within the PoC scope, Chapter 8 shows that: (i) the deterministic site core generates interpretable hourly demand and proportional dispatch trajectories across the committed epochs; (ii) the exported coupling artefact ΔE_t provides an auditable "thin-waist" linkage from the site pathway choice to a regional feasibility stressor at hourly resolution, and (iii) a regional proxy screening layer yields decision-relevant exposure summaries (exceedance and shedding exposure, discrete upgrade selection, and regret-style diagnostics). The evidence then extends to the 2035 site transition decision $p^{\text{site}} \in \{EB, BB\}$ evaluated under a paired-futures set Ω, while preserving the PoC boundary (no re-solving of dispatch and no feedback from the regional proxy to the site core).

Interpretation of all results is constrained by two guardrails that follow directly from the PoC design. First, the "regional layer" is a proxy screening module conditioned on an exogenous headroom signal H_t (as supplied via the signals interface) and a finite upgrade menu $\mathcal{P}^{\text{grid}}$. Second, the site dispatch kernel is proportional, providing face-valid operational narratives and closed accounting, but it is not a cost-minimising operational optimiser (which remains a deliberate next-fidelity extension). Consequently, the results should be read as screening-level evidence for interface exposure and robustness under uncertainty, sufficient to demonstrate auditability and decision relevance within the PoC scope, rather than as forecasts of system-optimal investment or operational.

To keep the chapter decision-focused, evaluation is organised around the following questions (claims under test):

- Deterministic site core plausibility: Does the site core generate plausible hourly demand trajectories and dispatch patterns by epoch that are interpretable and stable under the committed pathway backbone (DemandPack and unit-stack evidence)?

- Thin-waist coupling validity: Does the coupling artefact (ΔE_t) translate the site pathway choice into a regional feasibility stressor at hourly resolution, with timing and magnitude that can be audited against H_t?
- Regional screening under uncertainty: Under uncertain headroom and discrete upgrade options, how frequently does each pathway produce exceedance/shedding exposure, and which upgrade options are selected by the screening benchmark within $\mathcal{P}^{\text{grid}}$?
- Paired-futures decision robustness: Under paired futures Ω, what robustness indicators (e.g., failure rates, regret distributions, and win-rate style comparisons) discriminate EB versus BB without re-solving dispatch?
- Boundary coherence: Where do the results clearly support the PoC claims, and where do they indicate the need for the next fidelity step (e.g., cost-aware dispatch, richer feasibility-changing uncertainties, or iterative coordination)?

Table 8.1 provides the minimal provenance anchor for these experiments. For the bundle *poc_20260105_release02*, run identifiers *dispatch_prop_v2_capfix1*, epoch tags, pathway labels, the size of Ω, the headroom source used for screening (measured where available vs stylised proxy), and whether the post-processing overlay is applied.

Table 8-1 Artefact table (screening inputs)

Epoch	Pathway label	Layers executed	Futures (Ω)	Notes
2020	Committed	DP+SD+TP	–	Deterministic baseline for face-validity.
2025	Committed	DP+SD+ΔE+RS+TP	_	First electrification epoch; introduces ΔE/headroom comparison.
2028	Committed	DP+SD+ΔE+RS+TP	_	Higher electrification; stressor signal typically larger.
2035_EB	Option A	DP+SD+ΔE+RS+RDM+OVL+TP	100 paired	Grid RDM uses *Input/rdm/futures_2035.csv*
2035_BB	Option B	DP+SD+ΔE+RS+RDM+OVL+TP	100 paired	Same futures template for fairness

Layer codes: DP = DemandPack (hourly demand synthesis); SD = proportional site dispatch; ΔE= incremental electricity export; RS = headroom/tariff signals; RDM = grid screening + upgrades under futures; OVL = paired-futures robustness overlay; TP = *thesis_pack* outputs (curated figures/tables).

The input files, configuration and code that produced the results in this chapter are openly available (Process-Heat-RDM, tag thesis-poc-canonical, commit 368015d; Appendix E). Appendix D reproduces the baseline values so that the reported results can be read without access to the repository.

8.2 Deterministic site-core results across epochs

This section establishes the baseline credibility of the deterministic site core before any interpretation is drawn from exported coupling signals or regional screening outcomes. The PoC is intentionally structured around a committed pathway backbone across the historical and near-term epochs (2020, 2025, 2028), which provides a common, exogenous prefix for all subsequent analyses and ensures that inconsistent baseline modelling assumptions do not confound later decision comparisons in 2035. Accordingly, the evidence presented here is evaluated against face validity criteria: whether the resulting hourly demand and proportional dispatch trajectories are interpretable, stable under the committed assumptions, and internally consistent in accounting terms.

The key point is not to claim predictive accuracy, but to demonstrate that the core produces coherent temporal structure and sensible evolution across epochs (e.g., plausible seasonality, peak behaviour, and changes consistent with the declared electrification and production assumptions). This establishes a defensible foundation for the remainder of the chapter, where the exported incremental electricity artefact ΔE_t and the regional proxy screening layer are used to assess interface exposure and robustness under uncertainty.

8.3 Demand profile evolution (DemandPack outputs)

DemandPack is treated as an exogenous, contracted artefact in the PoC; accordingly, the reported demand profiles hold the intra-annual shape fixed across epoch tags and differ only through the declared annual target (and any peak-cap setting). The purpose of this subsection is therefore not to infer operational change, but to document artefact integrity and establish a stable baseline for interpreting downstream proportional dispatch outputs and the exported coupling artefact ΔE_t.

This subsection tests the first plausibility claim required for downstream analysis, that DemandPack produces stable, interpretable hourly heat-demand trajectories on a consistent hourly timebase, suitable for (i) auditable proportional dispatch and (ii) grid-facing screening using an hourly export trace. The DemandPack construction logic and interface gates are specified in Chapter 7; here, results are presented only to evidence face-validity and internal consistency (with the intention that future work can progressively calibrate, and ultimately replace, synthetic demand using metered operations where available).

Baseline year (2020): seasonal envelope, persistence, and peakiness

Seasonal structure and within-day spread. Figure 8.1 summarises the daily heat-demand envelope in 2020 using a median (P50) trace and a within-day spread band (P10–P90). The series exhibits a smooth seasonal envelope and a relatively consistent within-day spread, indicating that the generator produces persistent daily operating "regimes" rather than erratic hour-to-hour behaviour. This is the minimum behavioural coherence required before using the resulting hourly demand as the deterministic input to dispatch and export calculations.

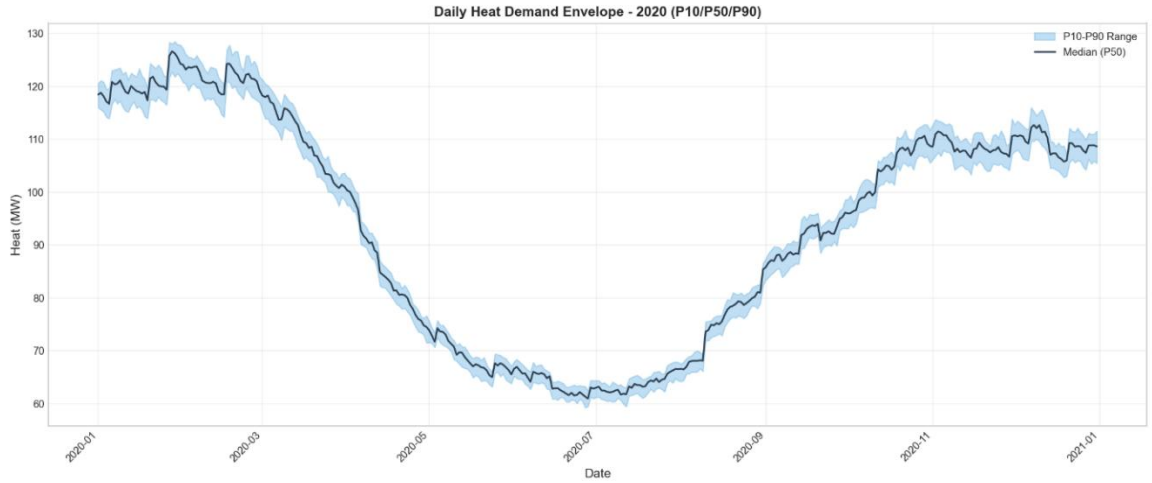


Figure 8.1 Daily heat-demand envelope in 2020 (MW)

Figure 8.2 shows the 2020 heat-demand load duration curve (LDC), constructed by sorting the full-year hourly heat demand series (8,760 hours) from highest to lowest; the x-axis therefore represents the cumulative share of annual hours rather than chronological time. Two properties matter for later coupling arguments: (i) a short upper tail (relatively few hours occur near the annual maximum), and (ii) a broad mid-range where "typical" demand is materially below the peak. This structure motivates why ΔE_t must be treated as timing-relevant: exceedance in the grid proxy depends on whether incremental load coincides with already-high demand hours, not solely on annual energy totals.

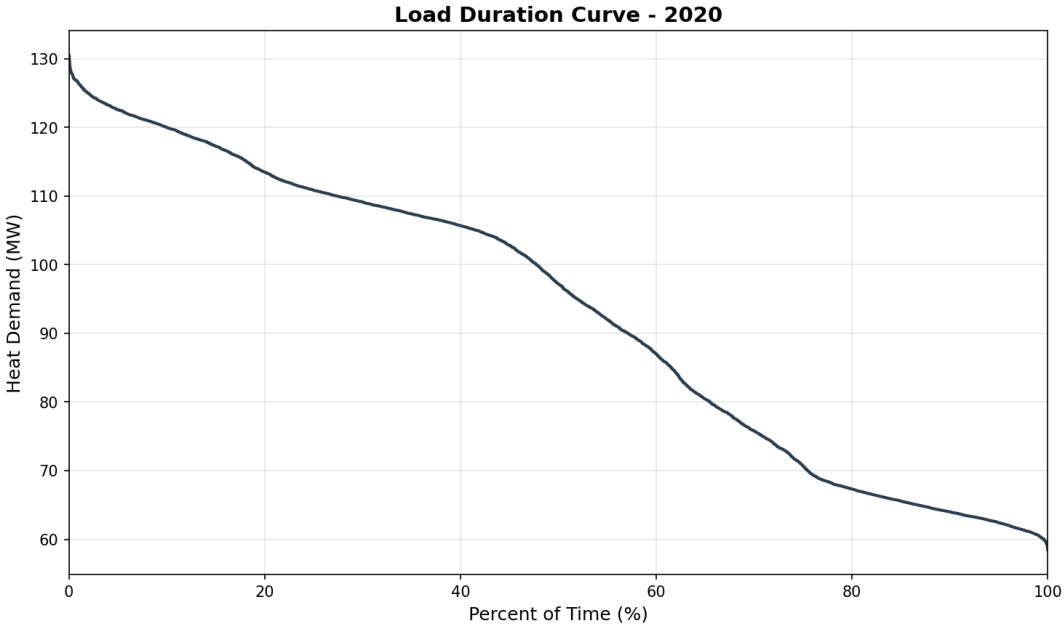


Figure 8.2 Load duration curve in 2020

Energy accounting and seasonal allocation. Figure 8.3 provides a monthly decomposition of total heat demand for 2020, which serves as an auditable check on annual energy closure and on how the seasonal envelope in Figure 8.1 aggregates over months. In this PoC run, the annual sum is 819.0 GWh, with clear seasonal allocation (e.g., January 89.4 GWh versus a mid-winter minimum of June 45.9 GWh, with July similarly low). The intent here is accounting transparency rather than interpretation of plant operations.

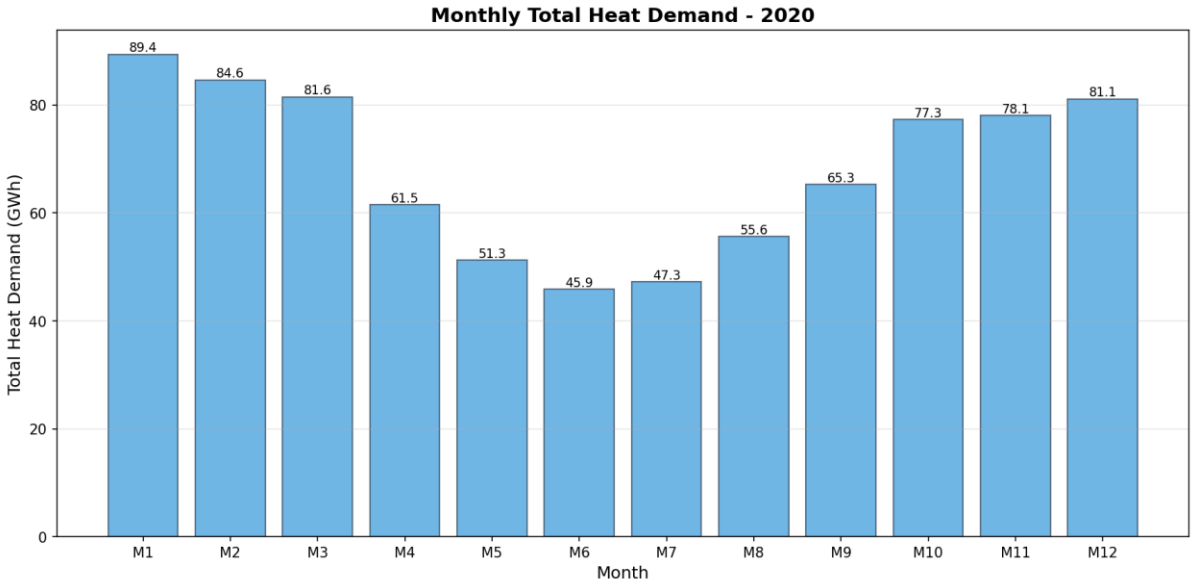


Figure 8.3 Monthly total heat demand in 2020 (GWh)

Intra-day and intra-week interpretability (diagnostics only)

To avoid repetitive narrative across epochs (given the fixed-shape contract), only one full diagnostic set is presented in the main text (2020); later epochs are treated as scaled variants under the same shape, unless and until additional defensible drivers are introduced (e.g., production-linked variability). Figures 8.4–8.5 therefore act as interpretability checks rather than additional empirical claims.

Figure 8.4 aggregates the mean diurnal profile by season in 2020. The profiles exhibit a coherent within-day structure (a broad plateau rather than sharp spikes) and a consistent separation across seasons, reinforcing that the synthetic generator is behaving as designed, a stable diurnal rhythm modulated by a seasonal envelope.

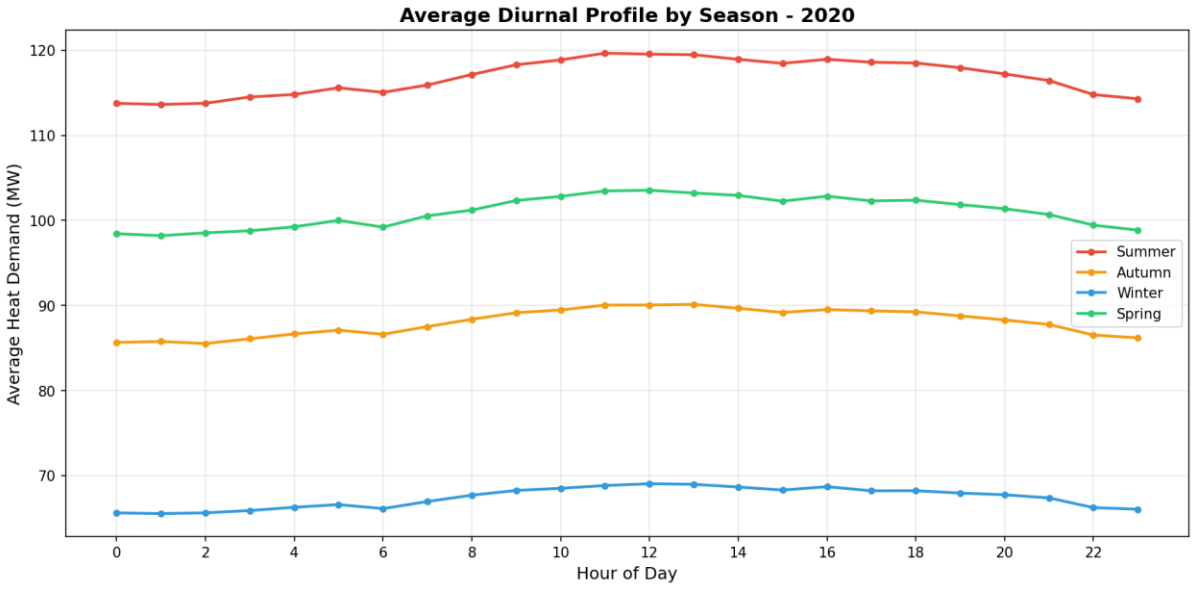


Figure 8.4 Mean diurnal profile by season in 2020 (MW)

Figure 8.5 shows an hourly trace for a representative week in February 2020 (a high-demand period). The week-scale chronology is smooth enough to remain operationally interpretable, while retaining sufficient intra-week variation to preserve meaningful hourly dynamics for subsequent dispatch and export artefacts.

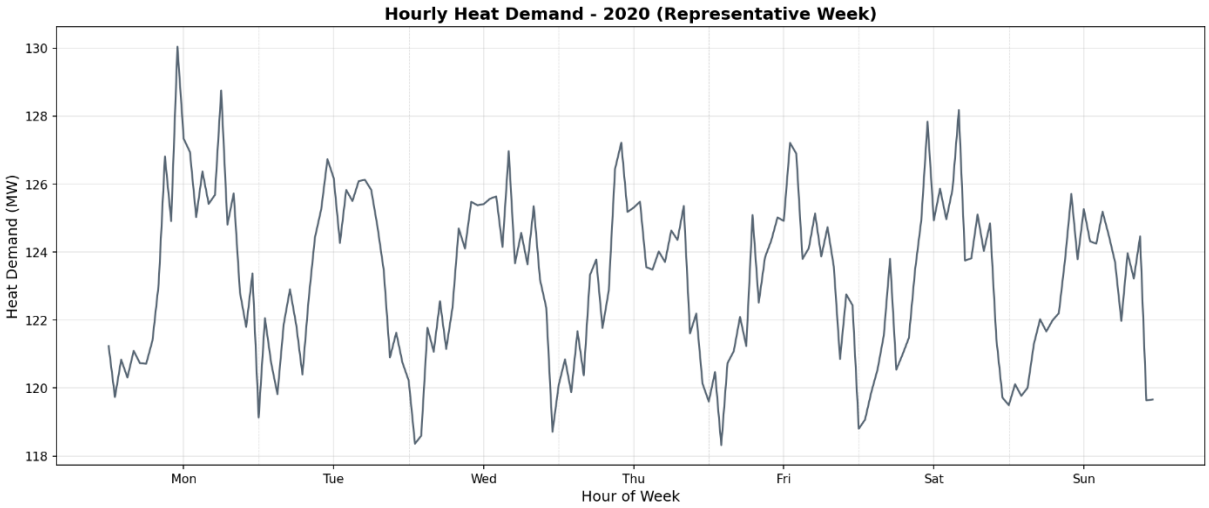


Figure 8.5 Hourly heat-demand trace for a representative week in February 2020 (MW)

Because demand is invariant (or scaled-only) across epochs by construction, any changes observed in later sections arise from the committed pathway backbone and the 2035 fork (EB versus BB), together with binding constraints in the grid screening layer rather than from unmodelled production variability. This provides a clean handoff to Section 8.3, which analyses the incremental electricity export trace and its interaction with hourly GXP headroom; Section 8.2.2 then reports how the fixed demand input is met by the pathway-specific utility portfolios under proportional dispatch.

For completeness, the corresponding DemandPack diagnostics for 2025, 2028, and 2035 are provided in Appendix B (Figures B.1–B.4) and confirm that the seasonal envelope, peakiness, and monthly allocation remain structurally consistent across epoch tags under the PoC demand-evolution assumption.

8.3.1 Proportional dispatch results across epochs

This subsection evaluates the second plausibility claim of the PoC: that the proportional site dispatch produces internally consistent and interpretable unit-level operating traces across epochs, and that the resulting annual summaries behave sensibly under the committed pathway narrative (coal → partial electrification → high electrification, with two alternative 2035 portfolios). The intent here is not to claim operational optimality, but to demonstrate that the deterministic site core can (i) allocate heat supply across a changing unit portfolio, (ii) respect declared maintenance availability, and (iii) yield stable, auditable annual accounting outputs that can subsequently underpin the exported coupling artefact ΔE_t and the regional screening layer.

A key property of the PoC site core is that feasibility is treated as an observable outcome rather than an implicit assumption. The dispatch ledger reports any residual heat requirement not met by the available portfolio as an explicit unserved term (energy and hours) For cases where unserved is zero, this provides a clean foundation for downstream interpretation: later differences in ΔE_t, headroom exceedance, or upgrade selection can be attributed to pathway composition and regional signals, rather than to hidden infeasibility in the site accounting. For cases where unserved is non-zero, it should be interpreted as a *stress indicator* conditioned on the declared portfolio and maintenance regime (decision-relevant evidence of operational fragility) rather than a numerical artefact. The purpose of the ledger is precisely to make shortfalls auditable rather than silently absorbed into balancing terms.

Table 8.2 summarises the primary ledger KPIs for proportional dispatch across epochs (from kpi_table_dispatch_prop_v2_capfix1, supplemented with unserved fields from the dispatch summary outputs). The table is used as evidence of coherent accounting and pathway-consistent evolution, not as a definitive economic comparison (given the simplified cost structure and the proportional dispatch kernel). Three points are emphasised. First, emissions fall sharply under the committed backbone as electrified heat displaces coal and reach zero in 2035_EB, while 2035_BB retains small residual emissions under the current parameterisation. Second, average heat cost rises across epochs under the PoC cost settings, reflecting assumed carrier costs and the increasing role of electrified supply; these values are carried forward primarily for regret-style diagnostics and relative comparisons under paired futures. Third, feasibility is confirmed (or stress is revealed) via the explicit unserved terms, not inferred indirectly from the penalty-cost field.

Table 8-2 Dispatch ledger summary (annual totals)

Epoch	Annual total cost (NZD m)	Avg. heat cost (NZD/MWh)	CO_2 (kt/yr)	Carbon cost (NZD m)	Unserved penalty (NZD)*
2020	14.2	17.3	244.7	7.3	≈ 0
2025	43.3	52.9	186.3	11.2	≈ 0
2028	83.5	102.0	65.9	4.6	≈ 0
2035_EB	91.0	111.1	0.00	0.00	≈ 0
2035_BB	76.8	93.8	1.6	0.1	≈ 0

* Values from the PoC dispatch summary outputs; penalties are ≈ 0 in the current run

Unit dispatch traces: portfolio evolution and operational interpretability

Figures 8.6–8.10 present the unit-level heat dispatch time series for each epoch. These plots serve two purposes. First, they provide a direct visual audit that the dispatch responds to the evolving portfolio composition (coal boilers in 2020; mixed coal–electric in 2025 and 2028; electric-dominant or biomass-supported portfolios in 2035). Second, they provide "shape checks" that the dispatch outputs remain interpretable under proportional allocation; for example, units contribute in proportion to availability and capacity in a way that produces stable, non-pathological trajectories suitable for downstream coupling.

In the baseline year (2020), dispatched heat is supplied exclusively by the four coal boilers (CB1–CB4), with the unit stack showing that multiple boilers contribute throughout the year rather than a single unit dominating supply (Figure 8.6). This provides a coherent baseline "portfolio fingerprint": coal units supply both the mid-load and upper-tail hours, consistent with a legacy configuration in which coal provides the full thermal envelope.

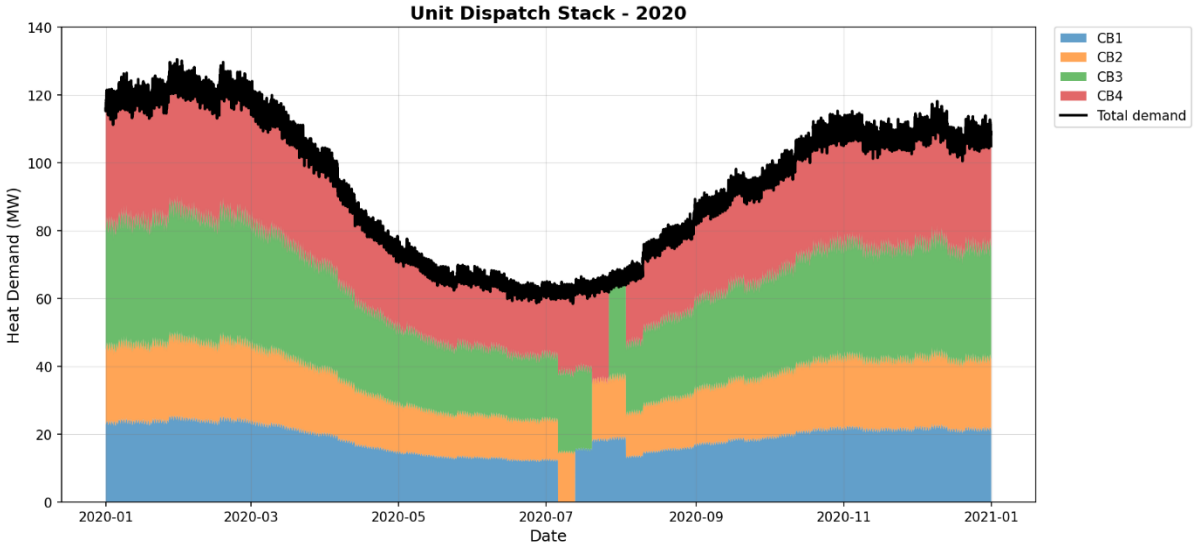


Figure 8.6 Unit-stack dispatch (2020)

In 2025, EB1 enters the portfolio and visibly displaces part of the coal contribution in the unit stack (Figure 8.7). The qualitative shift is twofold: (i) a new electrified heat tranche appears across the year, and (ii) the coal contribution contracts accordingly. This is the first epoch in which the site core begins to translate a pathway change into an electricity-linked operational implication, which becomes explicit in the exported ΔE_t series analysed in Section 8.4.

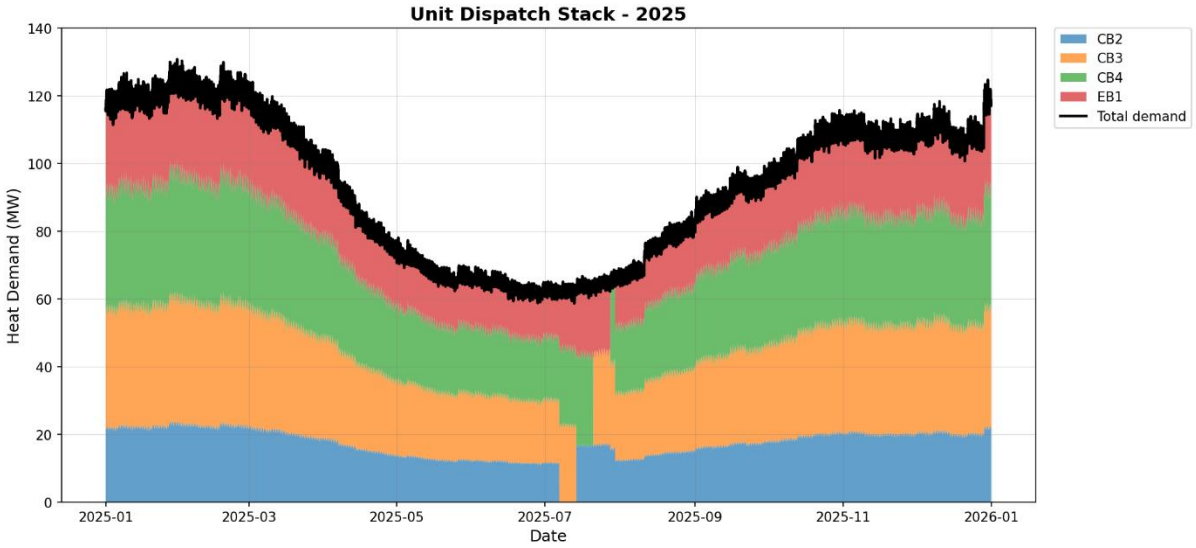


Figure 8.7 Unit-stack dispatch (2025)

By 2028, the unit stack shows that EB1–EB3 supply the majority of heat, with CB4 remaining as the residual thermal unit (Figure 8.8). This “transition” configuration is particularly useful as a face-valid intermediate state: a predominantly electrified portfolio with a small residual fossil component. It also foreshadows why ΔE_t must be treated as timing-relevant: as electrification increases, high-demand hours directly induce higher electrical import exposure.

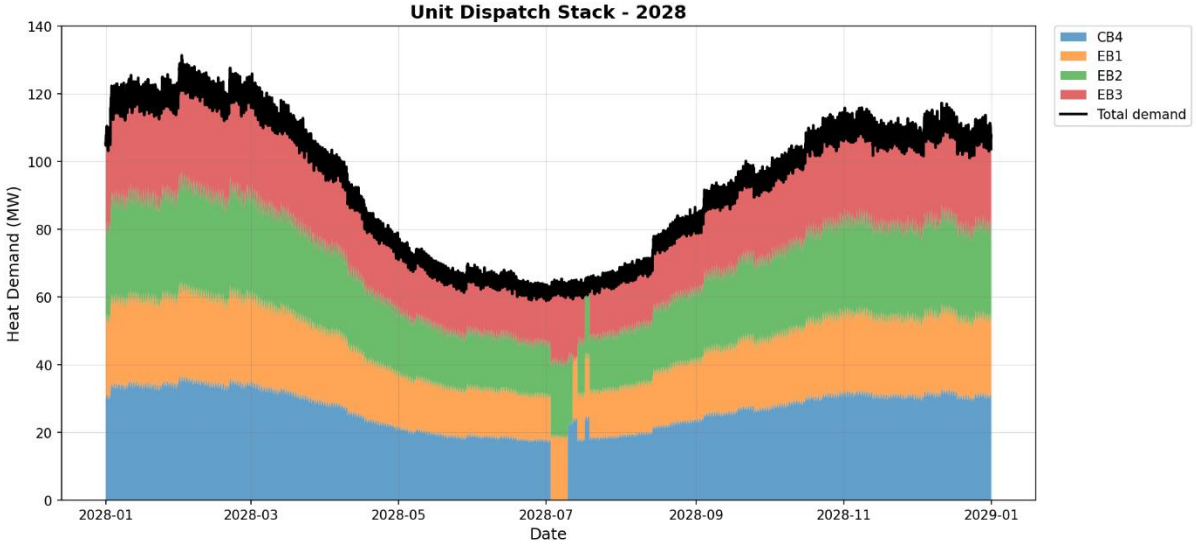


Figure 8.8 Unit-stack dispatch (2028)

The 2035 split is intentionally defined so that demand is held fixed (by the DemandPack contract) and the portfolio choice is the discriminating decision. The unit stacks provide the primary evidence that the two pathways implement distinct supply structures:

In the 2035_EB (fully electrified heat supply) pathway, heat is supplied by EB1–EB4 across the year, with no thermal fossil or biomass unit contributing (Figure 8.9). This provides the cleanest link to grid exposure in subsequent sections: the heat portfolio is fully electricity-coupled.

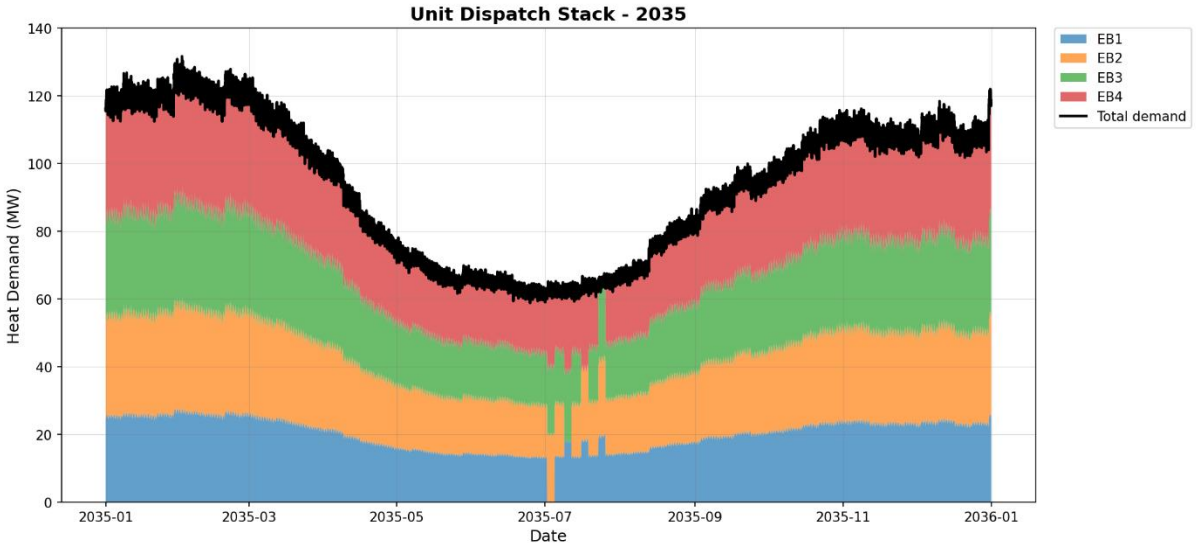


Figure 8.9 Unit-stack dispatch (2035_EB)

In the 2035_BB (biomass + electric hybrid) pathway, BB1 enters as a high-contribution thermal unit alongside EB1–EB3 (Figure 8.10). Relative to EB, the presence of BB1 provides an alternative heat supply channel that is not directly electricity-coupled, which is expected to reduce incremental electricity exposure (and therefore headroom pressure) in Section 8.4, all else equal.

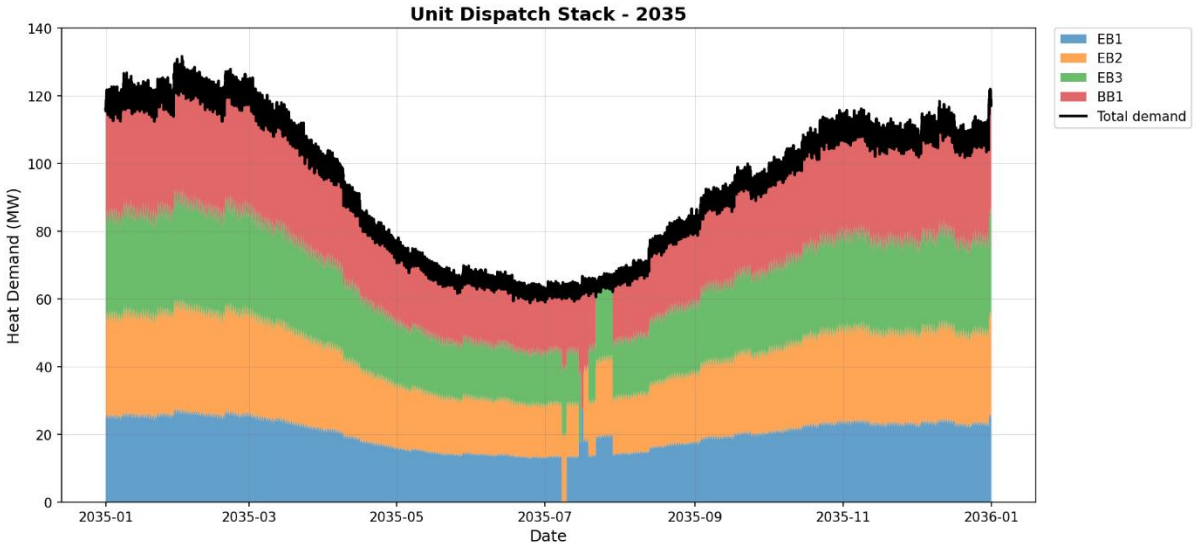


Figure 8.10 Unit-stack dispatch (2035_BB)

Utilisation duration curves (2035 only)

While the time-series plots show chronological behaviour, the utilisation duration curves provide a compact diagnostic of how intensively each unit is used across the year. Figures 8.11–8.12 present unit utilisation duration curves for the two 2035 pathways. Across epochs, the curves exhibit smooth, monotonic declines without discontinuities, supporting the interpretation that the dispatch is stable and proportional rather than driven by threshold effects or artefacts.

Two features are noteworthy:

- **Portfolio-dependent "top-end" utilisation:** The maximum output levels reflect the installed capacities and the high-load periods. In 2035_EB (Figure 8.11), the EB fleet collectively spans the full range of demand states, while in 2035_BB, BB1 assumes a comparable high-duty role alongside the electrified units, producing a compact fingerprint of the pathway choice. The fork alters which units carry the highest duty (electric-only versus hybrid electric and biomass).
- **Shape consistency and interpretability:** The broad similarity of curve shapes is consistent with proportional allocation, units tend to track one another through demand states, with divergence mainly attributable to capacity differences and maintenance availability. This is desirable for the PoC, because it yields predictable behaviour that can be traced cleanly into the exported coupling signal.

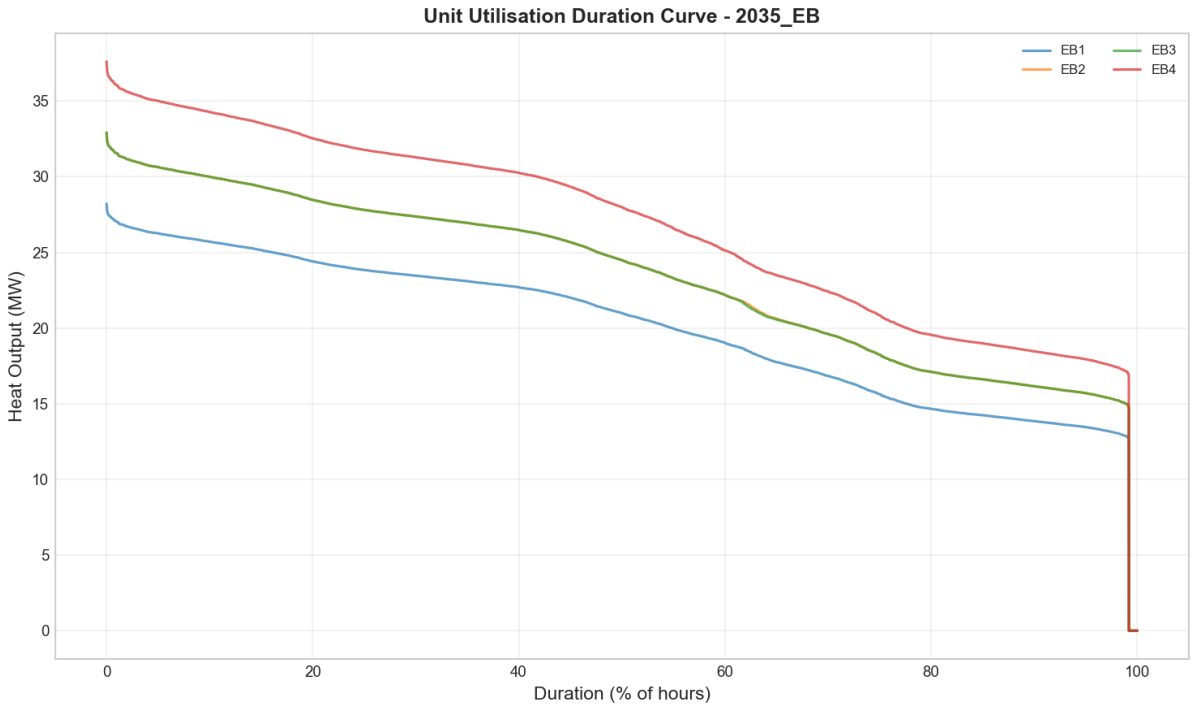


Figure 8.11 Unit utilisation duration curve (2035_EB)

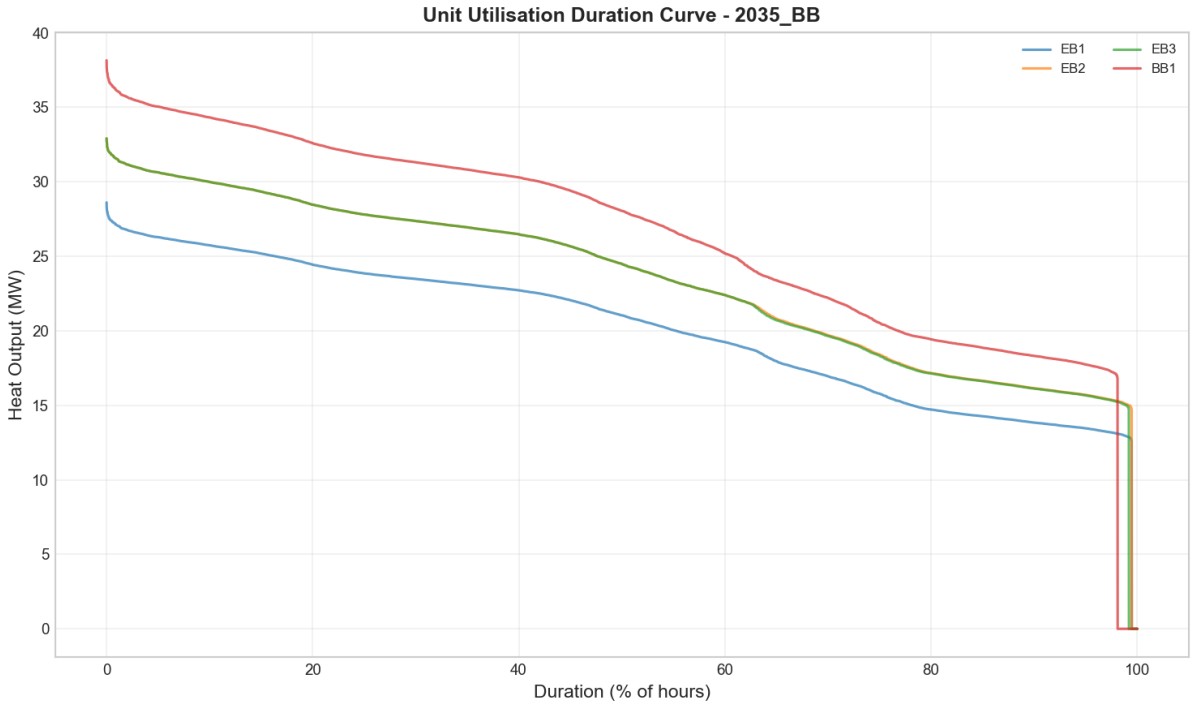


Figure 8.12 Unit utilisation duration curve (2035_BB)

Maintenance and availability conditioning: units-online evidence (2035 only)

Figures 8.13–8.14 report the proportion of units online under the declared maintenance windows for 2035_EB and 2035_BB. Both pathways experience mid-year availability reductions, but the BB case exhibits a deeper temporary reduction (falling to approximately half of units online at one point) relative to EB's predominantly three-of-four availability (Figures 8.13–8.14). These traces serve two roles in the chapter:

- They confirm that maintenance is applied as an availability constraint in the site core, where the dispatch is conditioned on time-varying unit availability rather than assuming full availability throughout.
- They identify periods in which coincident perturbations (e.g., demand growth, capacity derates, or concurrent grid constraints) would be most likely to tighten feasibility margins and amplify peak ΔE_t, precisely the conditions under which the downstream headroom screen is most likely to bind.

In addition, planned outages are treated as first-class exogenous constraints (availability matrix), and the resulting scarcity is reflected in visible changes in utilisation and (where relevant) explicit *UNSERVED*, rather than being hidden through capacity violations.

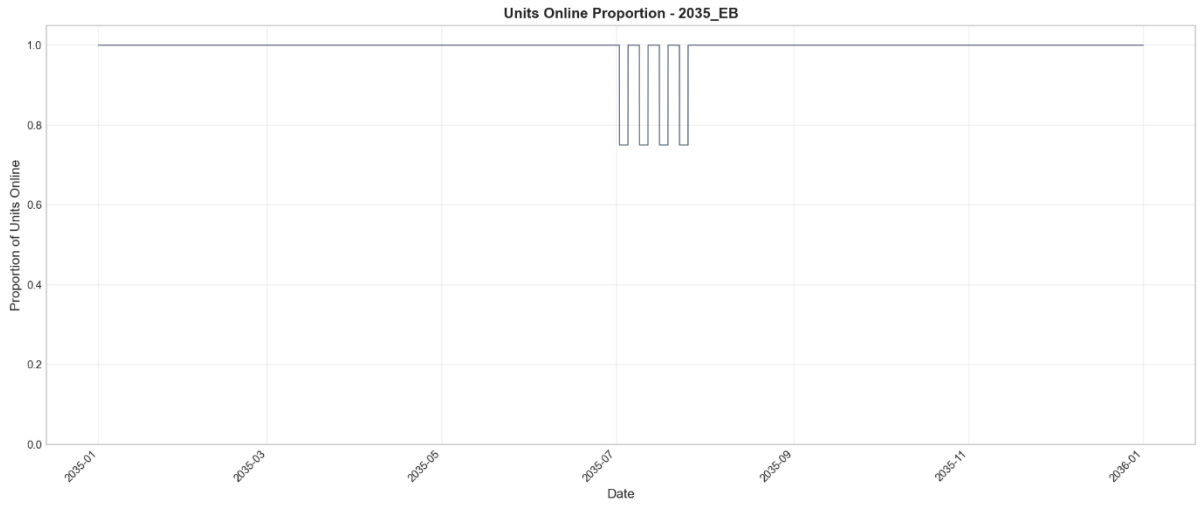


Figure 8.13 Units online proportion (2035_EB)

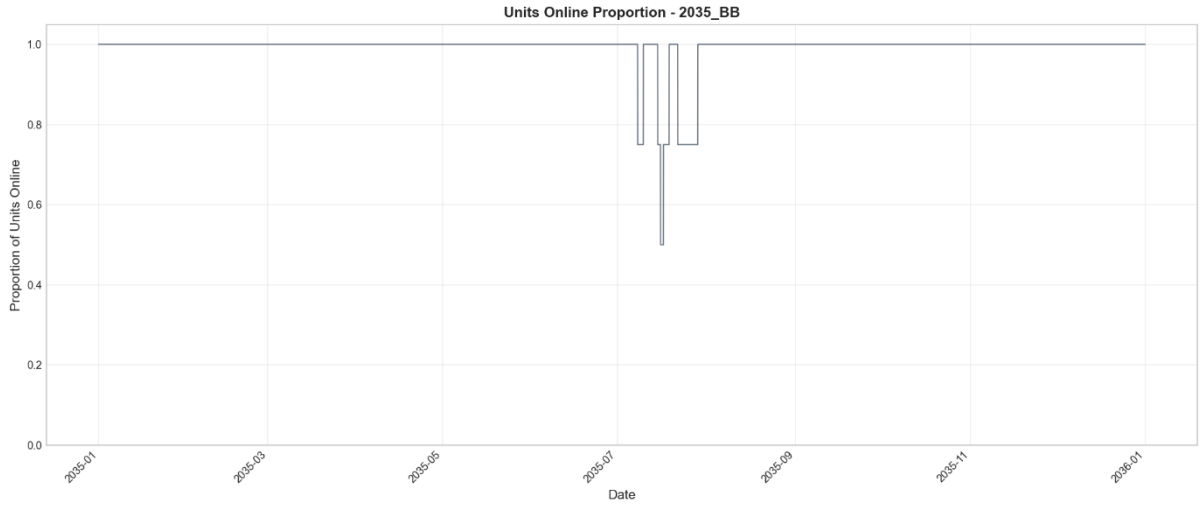


Figure 8.14 Units online proportion (2035_BB)

Together, the unit-stack plots (Figures 8.6–8.10), the 2035 duration-curve fingerprints (Figures 8.11–8.12), and the 2035 availability traces (Figures 8.13–8.14), anchored by the annual KPIs (Table 8.2), establish that the deterministic site core produces a coherent and auditable operating ledger across epochs. The committed pathway backbone yields the expected structural shift from coal to electrified heat by 2028, while the 2035 fork produces two clearly distinguishable portfolios (fully electric versus electric + biomass) under the same contracted demand. This provides the site-side preconditions and demonstrates that incremental electricity export is not an arbitrary signal, but a deterministic by-product of the pathway-specific dispatch ledger under declared availability, suitable for comparison against exogenous headroom in the regional screening layer.

8.4 Thin-waist coupling artefacts

This subsection evaluates the site exported coupling artefact $\Delta \mathrm{E}_{\mathrm{t}}$, and its interpretation relative to the exogenous GXP capacity and headroom signals used in the regional screening layer. The objective is not to claim predictive accuracy for Southland network operation (empirical GXP headroom traces are limited and the PoC uses a stylised headroom representation informed by available evidence), but to demonstrate that the framework produces a stable, auditable, hourly coupling artefact that (i) responds coherently to pathway choices at the site level and (ii) can be passed, without redesign, into the grid screening and RDM layer.

Concretely, $\Delta \mathrm{E}_{\mathrm{t}}$ is written as a thin-waist interface file at hourly resolution (e.g., *Output/signals/incremental_electricity_MW_<epoch_tag>.csv*). In the one-pass PoC, it is generated downstream of DemandPack and proportional dispatch and then consumed by the grid screening module as an incremental import requirement to be tested against the available headroom signal H_t. The analysis below therefore treats ΔE_t as the exported "coupling load" (site-side), and H_t as the exogenous "spare import margin" (grid-side). Their relationship establishes the plausibility of the coupling logic and motivates the robustness screen applied in later sections.

8.4.1 Exported incremental electricity trajectories

The exported series ΔE_t represents the additional grid electricity attributable to the site pathway and utility dispatch outcome, expressed relative to the baseline representation embedded in the PoC. In 2020, ΔE_t is set to zero by construction (coal-dominant baseline). From 2025 onwards, ΔE_t increases as electric heat supply enters the utility portfolio under the committed electrification narrative, and it diverges in 2035 depending on the decision pathway:

- 2035_EB: electrode-boiler pathway (higher electrification intensity → higher ΔE_t).
- 2035_BB: biomass pathway (additional heat supplied by biomass → ΔE_t remains closer to the committed-electrification level).

Importantly, the PoC currently treats non-utility, process-level electricity changes (e.g., process drives, refrigeration, compressed air) as out-of-scope for early epochs due to limited data separation. This is a scope choice, not a limitation of the interface; the same ΔE_t contract can be extended where metering or site disaggregation becomes available, so that ΔE_t consistently captures both utility-related and non-utility incremental electricity across all epochs.

Table 8.3 reports annual summaries computed directly from the exported hourly ΔE_t CSVs. The purpose is interpretability (orders of magnitude, step-changes across epochs, and pathway separation)

rather than calibration. Note that 2020 and 2028 contains 8,784 hours (leap year), while the other epochs contain 8,760 hours.

Table 8-3 Annual summary of exported incremental electricity (ΔE_t) across epochs

Epoch	Annual ΔE (GWh)	Mean (MW)	P50 (MW)	P95 (MW)	Peak (MW)
2020	0.00	0.00	0.00	0.00	0.00
2025	185.7	21.2	21.9	27.6	29.4
2028	593.1	67.5	70.6	88.0	94.8
2035_EB	827.3	94.4	98.9	123.8	133.0
2035_BB	592.9	67.7	70.7	88.4	95.0

The evolution of incremental electricity signals shows a clear, monotonic "electrification loading" signature across the committed transition epochs, and a clean pathway separation at the 2035 decision point. In the committed pathway, the step from 2020 to 2025 introduces a non-trivial but still moderate electric increment, consistent with an initial tranche of electrified heat supply entering the portfolio. the change from 2025 to 2028 is the dominant step: the annual exported ΔE_t increases by around 407 GWh and the mean level rises by approximately 46 MW, consistent with the narrative of additional electrode boilers commissioning and the consequent reallocation of utility heat supply towards electricity. This step-change is also expressed in the upper tail of the hourly distribution, the 95th percentile increases by about 60 MW and the peak hour increases by 65MW between 2025 and 2028, indicating that electrification is not only lifting the "base" incremental import but also materially increasing the coincidence risk of high incremental demand hours, precisely the feature that the thin-waist coupling artefact must preserve for downstream feasibility screening.

This shift in distributional shape is more informative than any single time-series trace. As illustrated by the exceedance (duration) curves across epochs in Figure 8.15, the 2025 curve is comparatively compressed, remaining below 40 MW throughout the year, whereas the 2028 curve occupies a higher band with a notably thicker upper tail (hours above 80 MW become a non-trivial fraction of the year). Importantly, the 2035_BB curve lies almost directly on top of the 2028 curve across most exceedance levels, evidencing that, under the PoC representation, switching the marginal 2035 option to biomass largely offsets what would otherwise be an additional electricity burden beyond the committed 2028 electrification level. This "near-overlay" is a useful internal-consistency check, that shows the exported coupling artefact responds in the expected direction to a pathway that deliberately substitutes electricity with an alternative heat carrier.

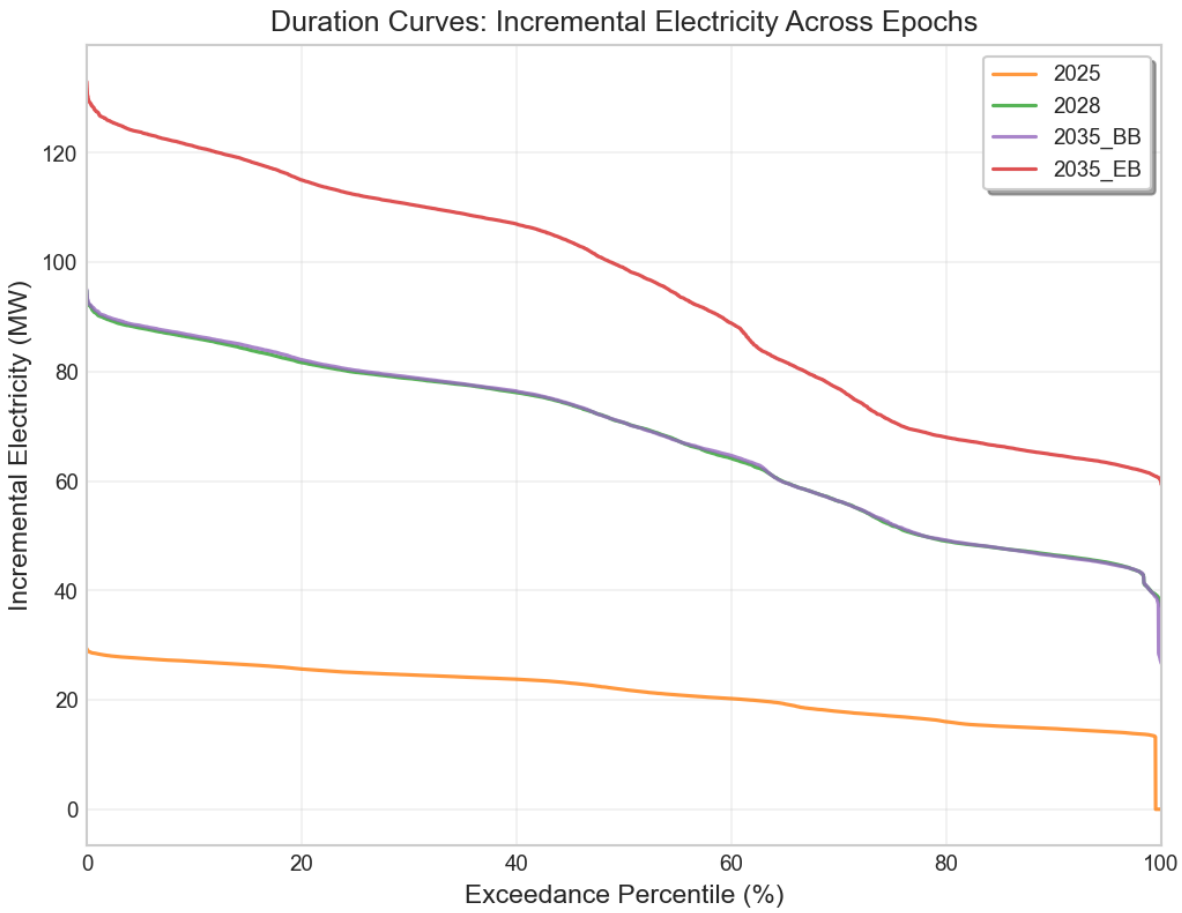


Figure 8.15 Duration curves of exported incremental electricity across epochs

The 2035 decision point then produces a clean and interpretable separation. Moving from 2035_BB to 2035_EB increases annual exported ΔE_t by around 234 GWh and raises the mean level by 27 MW, while also lifting the upper tail (P95 by approximately 35 MW and peak by 38 MW). In operational terms, the electrode-boiler pathway shifts the incremental import requirement into a regime where high incremental loads become common rather than exceptional (the duration curve shows that ΔE_t exceeds 100 MW for a substantial share of the year in 2035_EB, while remaining below 100 MW in 2028 and 2035_BB). The comparative summary bars in Figure 8.16 provide a noticeboard view of the same result: the committed electrification step dominates the transition from 2025 to 2028, and the 2035 decision then either preserves that committed level (biomass) or pushes the system into a materially higher incremental electricity regime (electrode boilers). Together, Table 8.3 and Figures 8.15–8.16 establish that the PoC produces an exported hourly coupling signal that is (i) monotonic with respect to committed electrification, (ii) pathway-separating at the decision epoch, and (iii) distributionally rich (capturing tail behaviour that will matter for grid feasibility), without requiring any redesign of the interface contract.

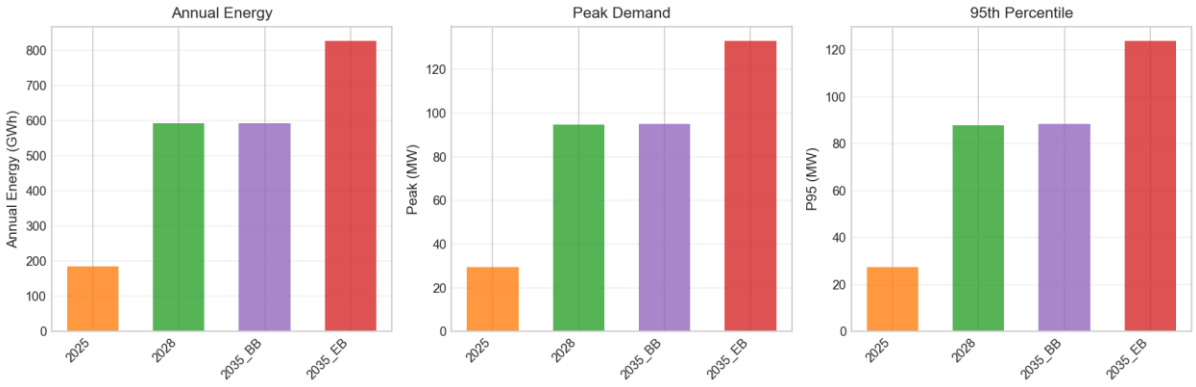


Figure 8.16 Summary comparison of exported incremental electricity metrics across epochs

These results should be interpreted as *capability evidence* rather than calibration. The absolute magnitudes reflect the PoC's stylised representation of utility electrification and the current scoping choice to exclude non-utility electricity evolution in early epochs; however, the *shape* and *ordering* properties demonstrated here are exactly what is required for a robust one-pass coupling workflow. In particular, because ΔE_t is exported as an auditable, hourly trace, it can be stress-tested against a range of plausible headroom trajectories even when empirical GXP headroom observations are limited, supporting the thesis claim that the framework can operate under partial observability by cleanly separating "site-side decisions" (which generate ΔE_t) from "grid-side feasibility signals" (which bound the admissible import). This motivates the next subsection, which treats headroom H_t as an interface feasibility diagnostic and examines the extent to which the rising ΔE_t tails in later epochs (and particularly under 2035_EB) create conditions under which headroom constraints are likely to bind.

8.4.2 Headroom feasibility diagnostics

In the PoC, the grid-facing feasibility signal is represented by an exogenous headroom trajectory H_t, interpreted as the available import margin at the Edendale GXP against which the exported coupling load ΔE_t is screened. This is deliberately not a claim about detailed Southland network operation: empirical headroom traces are limited, and the PoC adopts a stylised construction informed by available evidence to provide an auditable constraint signal at hourly resolution. Within the SignalsPack, the GXP connection capacity is held constants for all hours and all epochs, so binding behaviour in the PoC is driven by the residual margin H_t rather by variation in the nameplate capacity.

Across epochs, the headroom signal exhibits a coherent, interpretable annual structure. In the provided SignalsPack files, H_t spans approximately 30–75 MW, with a median close to 51–52 MW and a mean close to 51 MW. Interpreted against the fixed 200 MW connection limit, this implies baseline utilisation typically in the range of roughly 125–170 MW (i.e., $200 - H_t$), leaving a comparatively

modest margin for additional electrification-driven imports. The annual envelope is not dominated by short-lived extremes: a substantial fraction of the year sits in “tighter” headroom conditions (around the lower half of the distribution), with only episodic periods in the upper tail. This distributional picture is conveyed most directly by the headroom duration curve (Figure 8.17), which communicates the prevalence of constrained versus unconstrained conditions without requiring the reader to parse a full-year time series.

series exhibits a coherent and interpretable structure, it spans approximately 30–76 MW over the year, with a median close to 52 MW and a mean close to 51 MW. This implies that, under the PoC representation, baseline utilisation of the GXP is typically on the order of 125–170 MW (since H_t around 200 baseline import), leaving only a modest margin for additional electrification-driven imports. The seasonal component is a smooth annual envelope in which headroom is materially lower for a substantial share of the year (roughly half the hours at or below 50 MW), and only occasionally reaches the upper tail above around 70 MW. This structure is well conveyed by the duration curve shown in Figure 8.17, which communicate the distribution of “tight” versus “loose” headroom conditions without overloading the reader with full-year time series.

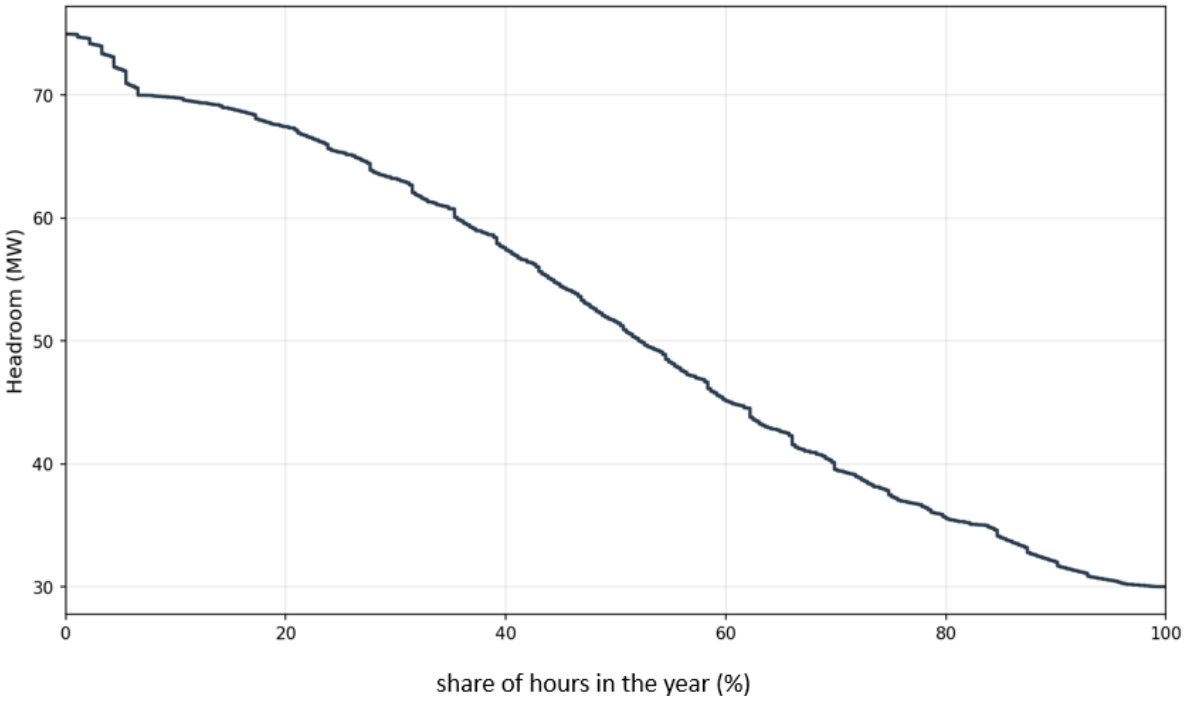


Figure 8.17 Edendale GXP headroom H_t duration curve (reference signal used for screening; 2035 headroom shown for illustration)

At intra-week scale, the headroom trajectory is intentionally “blocky”, exhibiting step changes rather than continuous operational variability (Figure 8.18). This reflects a deliberate modelling choice: in the

PoC, H_t functions as a capacity-like stress-testing surface rather than an operational power-flow result. The key requirement for the coupling demonstration is that the grid-side constraint is (i) transparent and reproducible, (ii) comparable across pathways and futures, and (iii) available at the same temporal resolution as the exported ΔE_t. Consistent with that requirement, the headroom series is identical between 2035 pathways (EB and BB) by construction, ensuring that any difference in feasibility outcomes at the decision fork is attributable to the site-side pathway and dispatch (via ΔE_t), not to inconsistent grid-side assumptions.

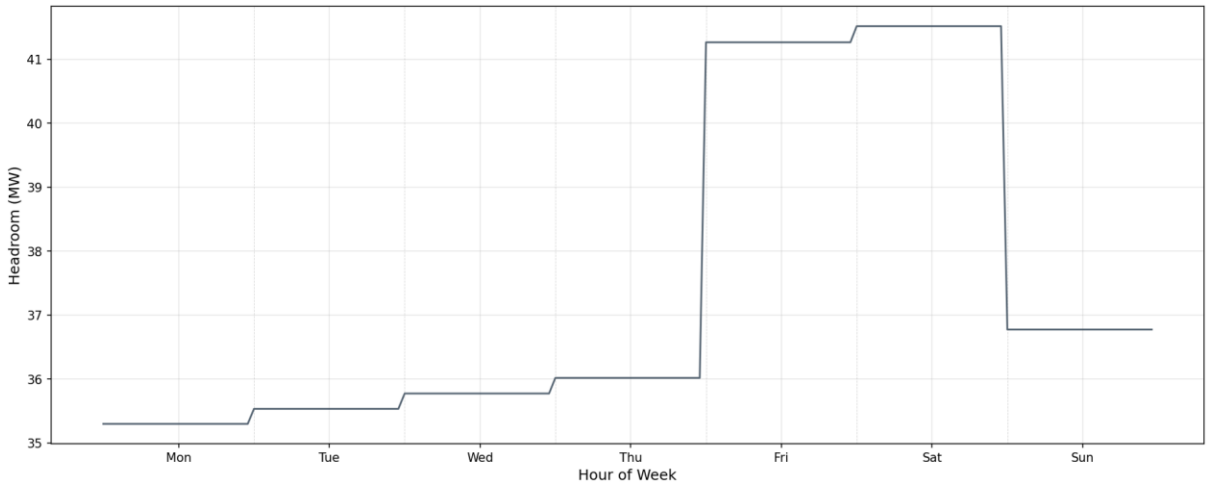


Figure 8.18 Representative-week time series of headroom H_t illustrating step-like intra-week structure (reference signal; 2035 shown)

Finally, for completeness of the thin-waist signal set, the SignalsPack also carries a flat energy tariff per epoch (constant for all timestamps within each epoch: 2020 = 105, 2025 = 145, 2028 = 130, 2035 = 110 NZD/MWh). In the PoC, this tariff is not used to generate H_t and does not affect feasibility screening directly; it exists to support transparent, reportable cost overlays once feasibility (and any exceedance handling) has been diagnosed using the ΔE_t–H_t interface.

8.4.3 Grid interface screening

With the headroom signal established, screening becomes a simple, auditable interface diagnostic. The hourly feasibility margin is $m_t = H_t - \Delta E_t$, and a binding hour occurs when $m_t < 0$ (equivalently $\Delta E_t > H_t$). For compact reporting, the magnitude of binding is expressed as the positive exceedance deficit

$$X_t \;=\; (\Delta E_t - H_t)^+ \;=\; \max(\Delta E_t - H_t,\, 0).$$

This deliberately reduces “grid feasibility” to a transparent guardrail that can be applied consistently across epochs and pathways, without embedding network physics inside the site model. The value of

the construct in the PoC is therefore not predictive accuracy but decision visibility: it makes explicit when committed electrification steps (and the 2035 pathway choice) push the exported coupling load into a regime where grid constraints would plausibly become decision-relevant.

In the early committed transition (2025), the screening outcome remains non-binding, ΔE_t stays below H_t throughout the year, and the tightest hours approach only a few megawatts of remaining margin. This is visible in the representative-week overlay (Figure 8.19), where the incremental import requirement tracks consistently beneath the headroom line. Interpreted as an interface result, 2025 therefore demonstrates that the framework can carry a non-trivial hourly coupling artefact without immediately forcing feasibility interventions, while still flagging that the system is operating closer to the headroom boundary than in the coal-dominant baseline.

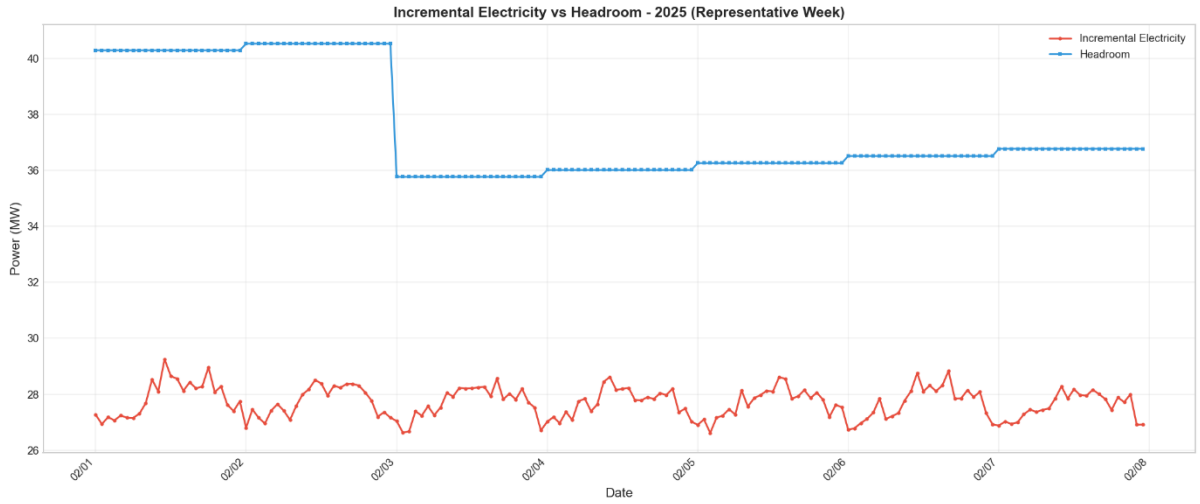


Figure 8.19 Representative-week overlay of incremental electricity ΔE_t and headroom H_t (2025)

The step from 2025 to 2028 moves the system into a qualitatively different regime. Under the stylised headroom signal, binding becomes frequent and structural: $\Delta E_t > H_t$ for approximately 64% of hours in 2028. The deficit X_t is typically on the order of tens of megawatts (median around 40 MW when exceedance occurs), with an upper tail approaching 60 MW in the most constrained hours. This regime shift is visually immediate in the 2028 overlay (Figure 8.20), where the incremental import trace sits persistently above the headroom line for most of the representative week. Interpreted as a coupling diagnostic, 2028 therefore marks the point at which a downstream robustness screen becomes a necessary complement to site-level electrification trajectories, rather than an optional refinement.

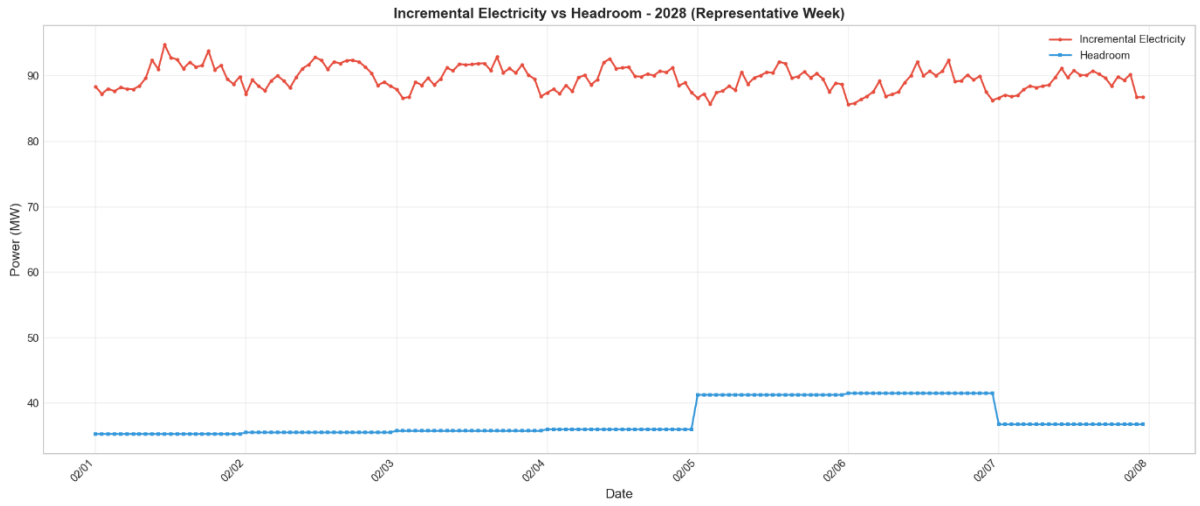


Figure 8.20 Representative-week overlay of ΔE_t and H_t (2028)

At the 2035 decision fork, the same screening logic produces a clean pathway separation while holding the grid-side signal fixed. The biomass pathway (2035_BB) largely preserves the 2028-like incremental electricity profile and therefore inherits a similar exceedance structure. The electrode-boiler pathway (2035_EB), by contrast, increases ΔE_t materially and pushes the system deeper into constraint: exceedance occurs for approximately 81% of hours, with a typical exceedance deficit on the order of 60–65 MW, and peak deficits approaching 100 MW in the tightest conditions. Figure 8.21 contrasts both pathways against the same headroom signal, making the pathway-driven nature of the feasibility divergence explicit.

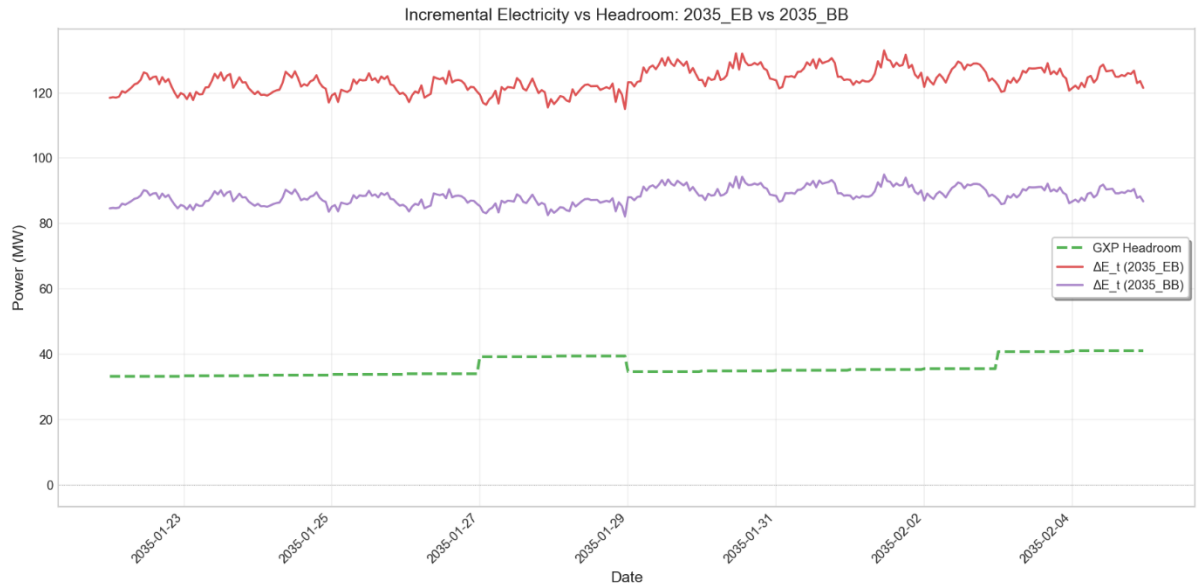


Figure 8.21 Representative-week overlay of ΔE_t and H_t at the 2035 fork (EB vs BB): pathway separation under a common reference headroom signal

To avoid over-interpreting a single representative week, Table 8.4 summarises full-year screening diagnostics computed directly from the hourly series. The table reports the distribution of H_t under the reference signal and the frequency and magnitude of binding hours via X_t. Two results are salient. First, 2025 remains non-binding (zero exceedance hours), whereas 2028 moves decisively into a binding regime with exceedance in roughly two-thirds of hours and peak deficits of order 60 MW. Second, at the 2035 decision fork the grid-side signal is held fixed, yet the exceedance regime separates cleanly by pathway: 2035_BB remains close to the 2028 committed-electrification case, while 2035_EB increases both the frequency and severity of deficits, roughly doubling the exceedance energy proxy under the reference headroom signal. This is the behaviour required for a decision-facing robustness screen: the exported coupling artefact ΔE_t yields materially different interface stress profiles under a discrete site decision, without redesigning interfaces or introducing second-pass feedback.

Table 8-4 Reference headroom screening diagnostics based on hourly exceedance deficit $X_t = (\Delta E_t - H_t)^+$

Epoch	H_t P50 (MW)	H_t P95 (MW)	Exceedance hours (%)	Mean deficit (MW)	Peak deficit (MW)	Deficit energy (GWh)
2020	51.3	71.8	0.0	0.0	0.0	0.0
2025	51.4	71.8	0.0	0.0	0.0	0.0
2028	51.3	71.9	63.6	35.2	59.5	196.7
2035_BB	51.6	72.1	64.0	35.2	60.3	197.4
2035_EB	51.6	72.1	80.6	54.7	97.7	386.2

Notes:

(i) 2020 and 2028 contain 8,784 hours (leap years); 2025 and 2035 contain 8,760 hours.

(ii) "Deficit energy" is the annual sum of exceedance deficits (MWh) reported in GWh; it is a screening-level stress indicator under the reference headroom signal.

These screening results should not be read as forecasts of actual Southland constraints. Their role in the PoC is to demonstrate the intended capability of the framework: (i) the site module produces a stable, auditable hourly coupling artefact whose magnitude and distribution respond coherently to committed electrification steps and to a discrete pathway choice, and (ii) the exported artefact can be screened against a grid-side feasibility envelope in a way that naturally motivates the subsequent robustness treatment (upgrade-versus-shed logic under deep uncertainty) without introducing a second-pass feedback loop.

8.5 Grid robustness screen (RDM) results across futures

This section reports the results of the grid robustness screen introduced in Chapter 7. A one-pass, post-dispatch evaluation in which the site's exported incremental electricity profile is tested against uncertain GXP headroom and a discrete set of candidate upgrade policies. The objective is not to reproduce a full power-system planning model, but to demonstrate the decision-analytic capability of the proposed modular framework to (i) represent deep uncertainty at the site–grid interface (headroom availability, upgrade cost, and outage valuation), (ii) evaluate discrete infrastructure options as policies, and (iii) summarise performance using ensemble-based robustness measures (satisficing rates and regret).

All monetary results in this section are grid-interface costs only, including the annualised cost of any selected capacity uplift plus penalties associated with unserved energy (and any consent or overrun uplifts). Upgrade costs are reported as epoch-scoped screening requirements, not as incremental costs conditional on an explicit prior-upgrade state. They should therefore be interpreted as the marginal grid consequence of a given site pathway, reported here to demonstrate how the framework can host robust evaluation in a separate module via artefact exchange, without entangling site optimisation and regional screening.

The robustness screen is evaluated over 100 paired futures, shared across 2035_BB and 2035_EB, so that pathway comparisons later (Section 8.5) are conditioned on an identical uncertainty ensemble. Each future perturbs a small set of uncertain drivers using multiplicative factors (e.g., available headroom and incremental-load scaling). The upgrade policy set comprises:

- **S0:** No upgrade (allow shedding)
- **S1:** RETA-style incremental upgrade option 1 (+21 MW, N)
- **S2:** RETA-style incremental upgrade option 2 (+32 MW, N–1)
- **S3:** RETA-style incremental upgrade option 3 (+97 MW, N)
- **S4:** PoC extension beyond RETA, representing a larger incremental upgrade option 4 (+150 MW, N)
- **AUTO:** ex post oracle benchmark, which selects the least-cost option among the candidate set (used for regret and for illustrating switching pressure, not as an implementable "foresight" policy).

8.5.1 Strategy dominance and robustness under 2035_BB

Under the 2035_BB pathway, the grid robustness screen identifies a clear dominance structure. Low-capacity options (S0–S2) perform poorly across the ensemble, producing persistent unserved energy and correspondingly high penalty exposure. In contrast, S3 eliminates unserved energy across all sampled futures while avoiding the higher annualised uplift cost of S4. This is evident in the AUTO selection response, where across the full range of headroom realisations, AUTO selects the same uplift level (corresponding to S3) in 100% of futures (Figure 8.22). This indicates that, for 2035_BB, the uncertainty ensemble does not generate conditions under which the larger uplift (S4) produces incremental reliability value.

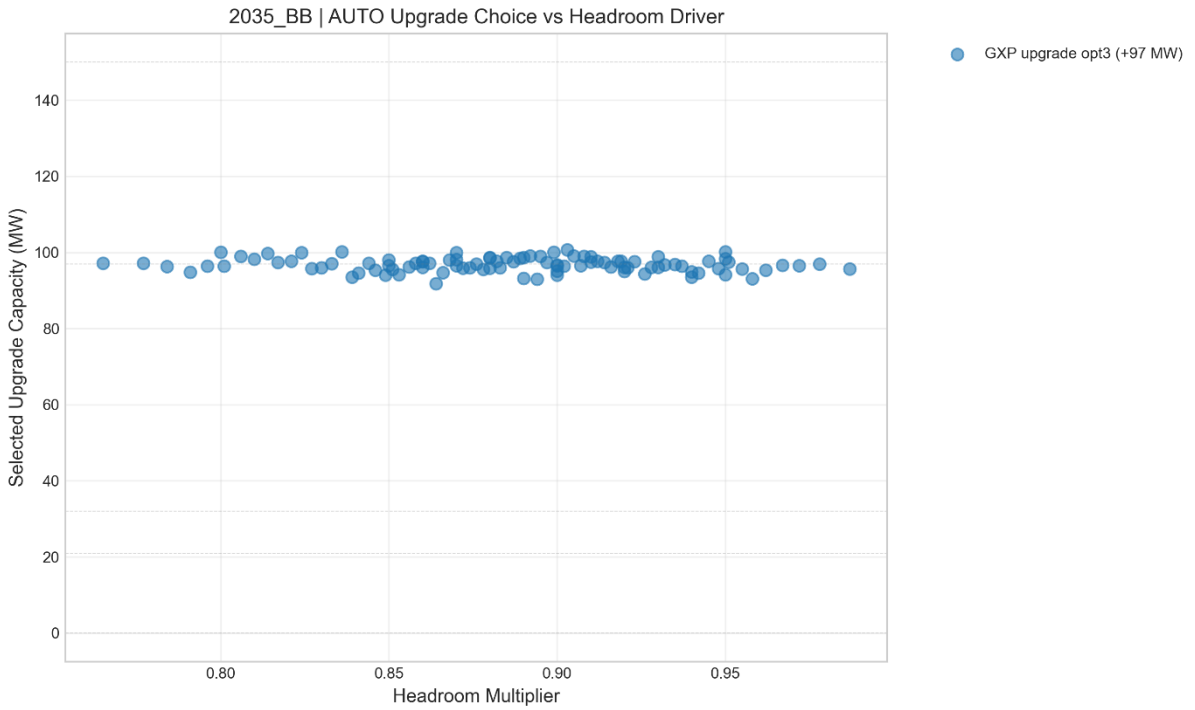


Figure 8.22 Auto-selected grid upgrade choice vs headroom across 100 futures (2035_BB)

The annual shedding distributions reinforce the same conclusion. As illustrated in Figure 8.23, strategies S3, S4 and AUTO deliver zero unserved energy across all sampled futures; consequently, on a logarithmic y-axis the corresponding boxplots collapse at the plot floor (exact zeros cannot be displayed on a log scale and therefore appear as a tight cluster at the axis minimum). By contrast, S0, S1 and S2 exhibit materially non-zero annual shortfalls across the ensemble, with the box/whisker structure capturing cross-future variation driven by uncertain headroom and incremental-load scaling (rather than any stochasticity within the dispatch). Interpreted as a robustness diagnostic, Figure 8.23 therefore separates “adequacy” (S3+) from “structural shortfall” (S0–S2) in a visually unambiguous

way. S3, S4 and AUTO have zero shed in every future and therefore do not appear on the logarithmic axis.

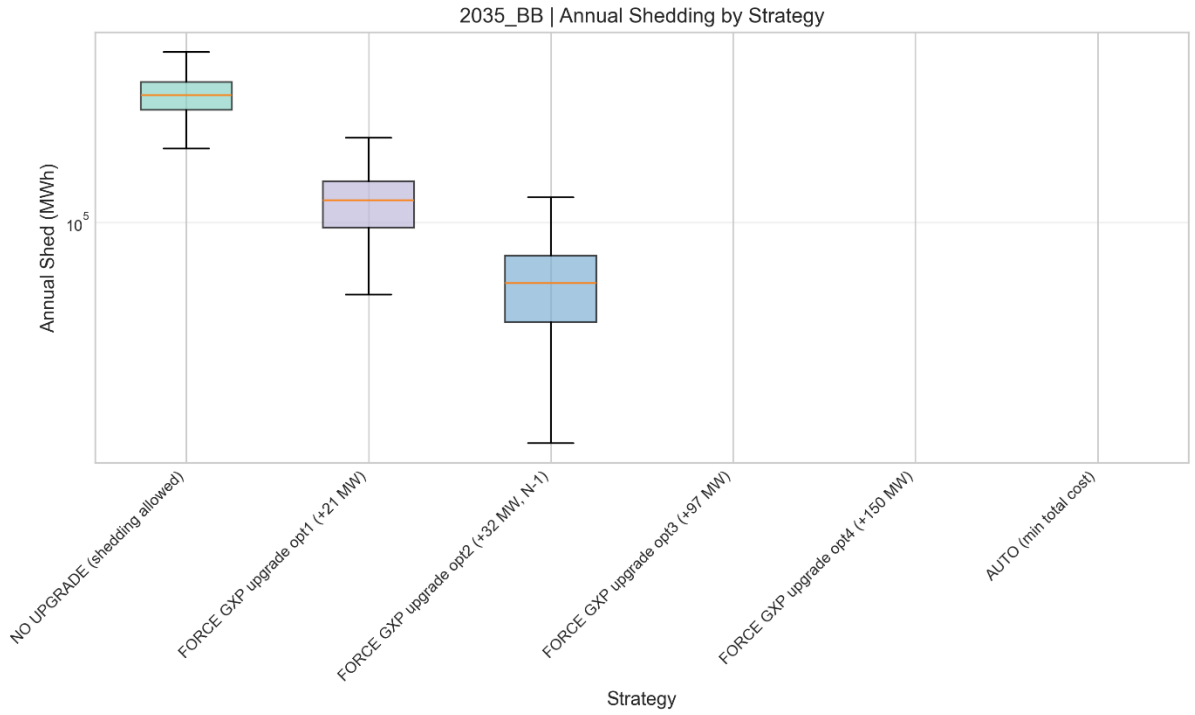


Figure 8.23 Annual shed by strategy across futures (2035_BB)

Table 8.5 quantifies these findings. Using a strict reliability satisficing rule (*annual_shed* = 0), S3 achieves a 100% no-shed rate and has low, tightly bounded grid-interface costs driven primarily by uncertainty in upgrade cost multipliers. By contrast, S0–S2 are uniformly non-satisficing and incur very large penalty-heavy costs under the assumed VOLL settings.

The very large grid-interface costs observed for low-upgrade strategies are dominated by the unserved-energy penalty, not by capital expenditure. In the screening model, total grid-interface cost is computed as the annualised cost of the selected capacity uplift plus a value-of-lost-load (VOLL) penalty applied to unserved energy. VOLL is intentionally high (here spanning 5,000–20,000 NZD/MWh across the sampled futures), so that strategies that rely on material shedding are treated as decision-inferior even if they avoid upgrade expenditure. Consequently, costs in the multi-billion NZD per year range should be read as a reliability stress score (a screening penalty), not as a forecast of actual cash outlays. What is decision-relevant is the dominance ranking and the tail-risk structure (how often and under what futures shedding persists), rather than the absolute magnitude of penalised outcomes.

Table 8-5 Grid robustness summary by strategy across 100 futures (2035_BB)

Strategy	ΔCap (MW)	AUTO select (%)	No-shed (%)	Median cost (m NZD)	P90 cost (m NZD)	Max shed (GWh)	Max regret (m NZD)
S0	0	0	0	2920.12	4557.17	302.381	6042.24
S1	21	0	0	1405.18	2477.17	173.207	3443.83
S2	32	0	0	836.79	1544.49	117.491	2324.46
S3	97	100	100	5.54	7.06	0.000	0.00
S4	150	0	100	7.43	9.47	0.000	3.06
AUTO	—	—	100	5.54	7.06	0.000	0.00

Note: Costs are grid-interface only (annualised uplift + unserved penalty + consent uplift). “Regret” is computed relative to the ex-post best strategy per future; AUTO has zero regret by construction.

Consequently, for 2035_BB, the robustness screen yields an unambiguous decision signal, a moderate uplift (S3) is sufficient to render the pathway grid-feasible across the tested futures, and further uplift is weakly dominated on cost grounds. This is precisely the kind of output the screening layer is designed to produce, a defensible “infrastructure adequacy” recommendation conditional on a site pathway, rather than a fragile single-scenario feasibility statement.

8.5.2 Strategy dominance, tail risk, and switching behaviour under 2035_EB

The 2035_EB pathway exhibits qualitatively different behaviour. Low-upgrade options (S0–S2) remain clearly non-robust, with persistent unserved energy and very high penalty costs. However, the comparison between the two higher upgrades becomes decision-relevant: S3 substantially reduces unserved energy but does not eliminate it in all futures, whereas S4 removes unserved energy across the ensemble at the cost of a higher annualised uplift charge.

This manifests as switching in the AUTO benchmark. Figure 8.24 shows that AUTO selects S3 (around 97 MW uplift) in 51% of futures and S4 (around 150 MW uplift) in 49% of futures. Importantly, this figure links the choice directly to headroom uncertainty, where unlike 2035_BB, electrification introduces a region in the uncertainty space where the larger uplift becomes cost-justified.

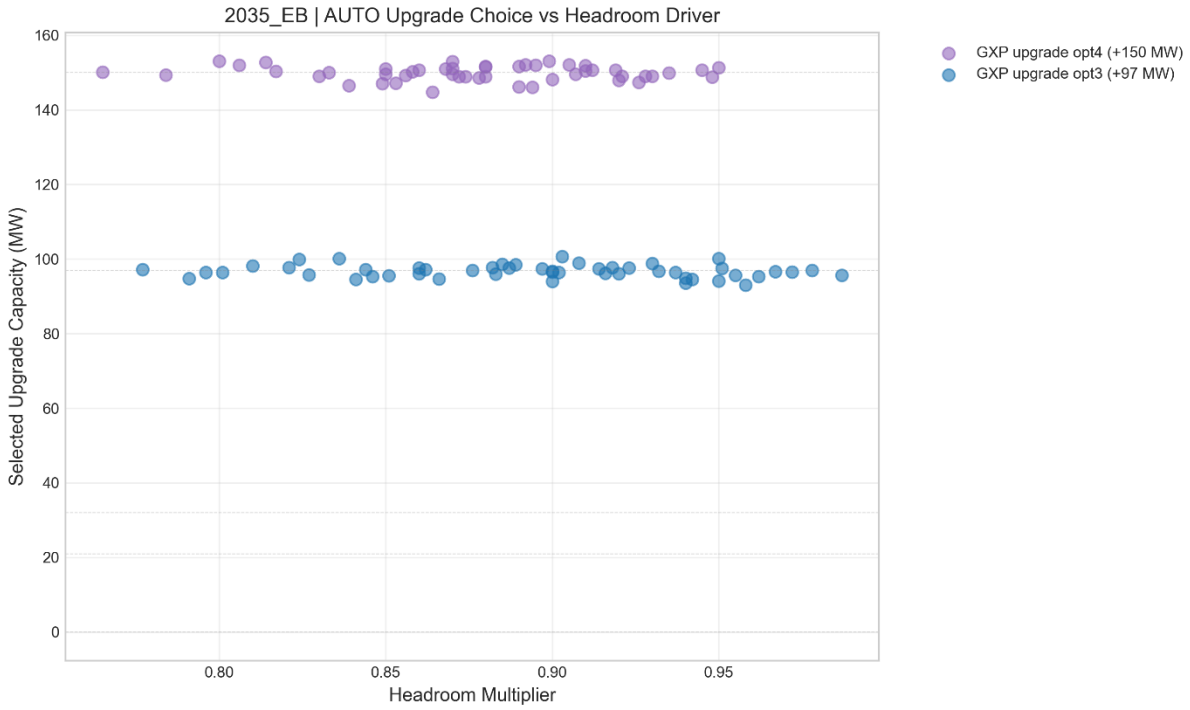


Figure 8.24 Auto-selected grid upgrade choice vs headroom across 100 futures (2035_EB)

The distribution of annual unserved energy clarifies the upgrade trade-off under 2035_EB. Strategy S4 eliminates shedding across all futures, whereas S3 retains a non-zero upper tail of annual shortfall in adverse futures (even though typical deficits remain far smaller than under S0–S2). Figure 8.25 overlays individual future realisations (markers) on the box/whisker summaries, and because the y-axis is logarithmic, futures with exactly zero shedding appear as a cluster at the axis floor; this should be read as “no shortfall”, not as an artefact.

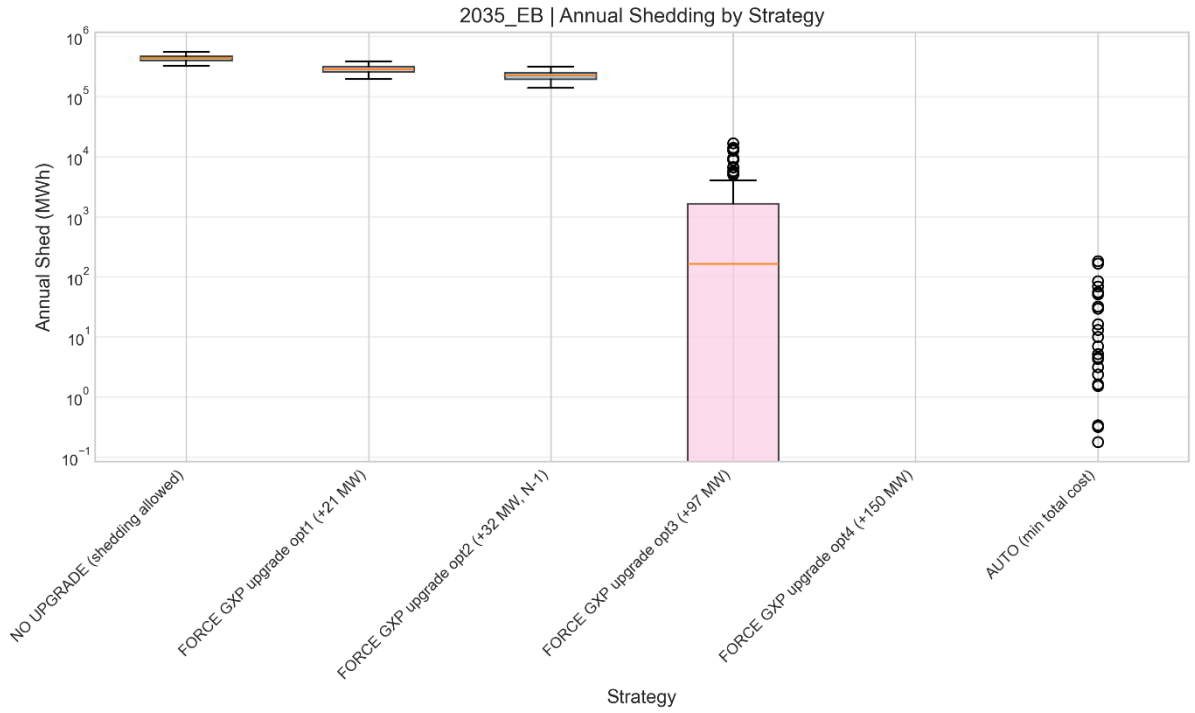


Figure 8.25 Annual shed by strategy across futures (2035_EB)

Figure 8.26 visualises the ensemble trade-off between reliability and cost by plotting, for each strategy and future, total grid-interface cost against annual unserved energy (each point corresponds to one future under one strategy). The pattern is strongly structured: strategies with insufficient uplift (S0–S2) cluster in the high-shed region and incur very large costs because the VOLL penalty dominates whenever shedding is material. By contrast, S4 collapses into the near-zero shedding region across the full ensemble, converting tail-risk into a predictable annualised uplift cost. Strategy S3 occupies an intermediate position: it is frequently close to the efficient frontier, but it retains a non-zero upper tail of shedding under adverse futures, which becomes economically salient when incremental-load scaling is high and headroom is degraded. AUTO traces the lower envelope by selecting the least-cost option in each future; its switching between S3 and S4 therefore reflects the point at which the marginal uplift cost of S4 is outweighed by avoided tail penalties. Figure 8.26 thus makes the dominance mechanism explicit: the “expensive” strategies are expensive because they represent penalised reliability failure, not because upgrade capex is intrinsically large.

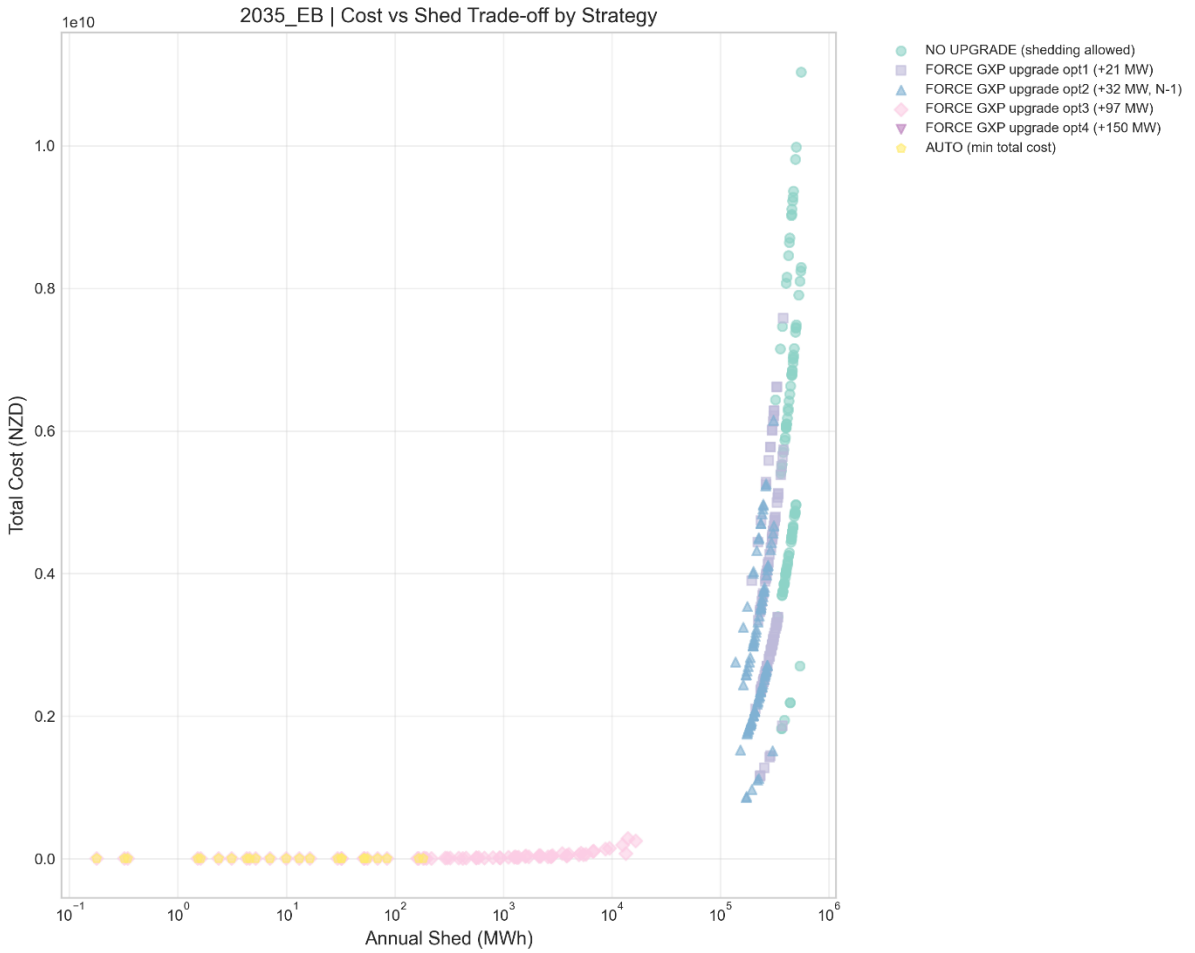


Figure 8.26 Cost–shed trade-off scatter by strategy (2035_EB); cost axis in NZD ($\times 10^{10}$); shed on a logarithmic axis

Table 8.6 provides the corresponding robustness metrics: S3 achieves high (but not complete) no-shedding performance, whereas S4 yields full reliability within the tested futures and reduces tail costs associated with penalised shortfalls.

Table 8-6 Grid robustness summary by strategy across 100 futures (2035_EB)

Strategy	ΔCap (MW)	AUTO select (%)	No-shed (%)	Median cost (m NZD)	P90 cost (m NZD)	Max shed (GWh)	Max regret (m NZD)
S0	0	0	0	5617.56	8650.31	552.940	11028.58
S1	21	0	0	3571.40	5772.28	382.293	7572.84
S2	32	0	0	2712.33	4576.48	311.220	6142.37
S3	97	51	27	7.47	66.21	16.557	279.38
S4	150	49	100	7.43	9.47	0.000	2.86
AUTO	—	—	76	6.75	8.81	0.182	0.00

A scenario-discovery style diagnostic is provided in Figure 8.27, which can be read as a data-driven vulnerability map for the AUTO (ex post) benchmark. Rather than treating uncertainty as a vast space to be exhaustively explored, the heatmap highlights the subset of futures in which the least-cost

response still involves non-zero residual shortfall (i.e., AUTO chooses a lower-uplift option and accepts a small amount of shedding rather than paying the additional annualised uplift cost). In the present ensemble, AUTO exhibits non-zero shedding in 24 of 100 futures (Table 8.6), concentrated in a relatively narrow region characterised by headroom multipliers of approximately 0.79–0.99 and incremental-load multipliers of approximately 0.94–1.02. Importantly, the magnitude of this residual is small (AUTO max shed is around 0.182 GWh; Table 8.6), so the heatmap should be interpreted as identifying where the decision boundary lies, not where the system "fails" in a material sense.

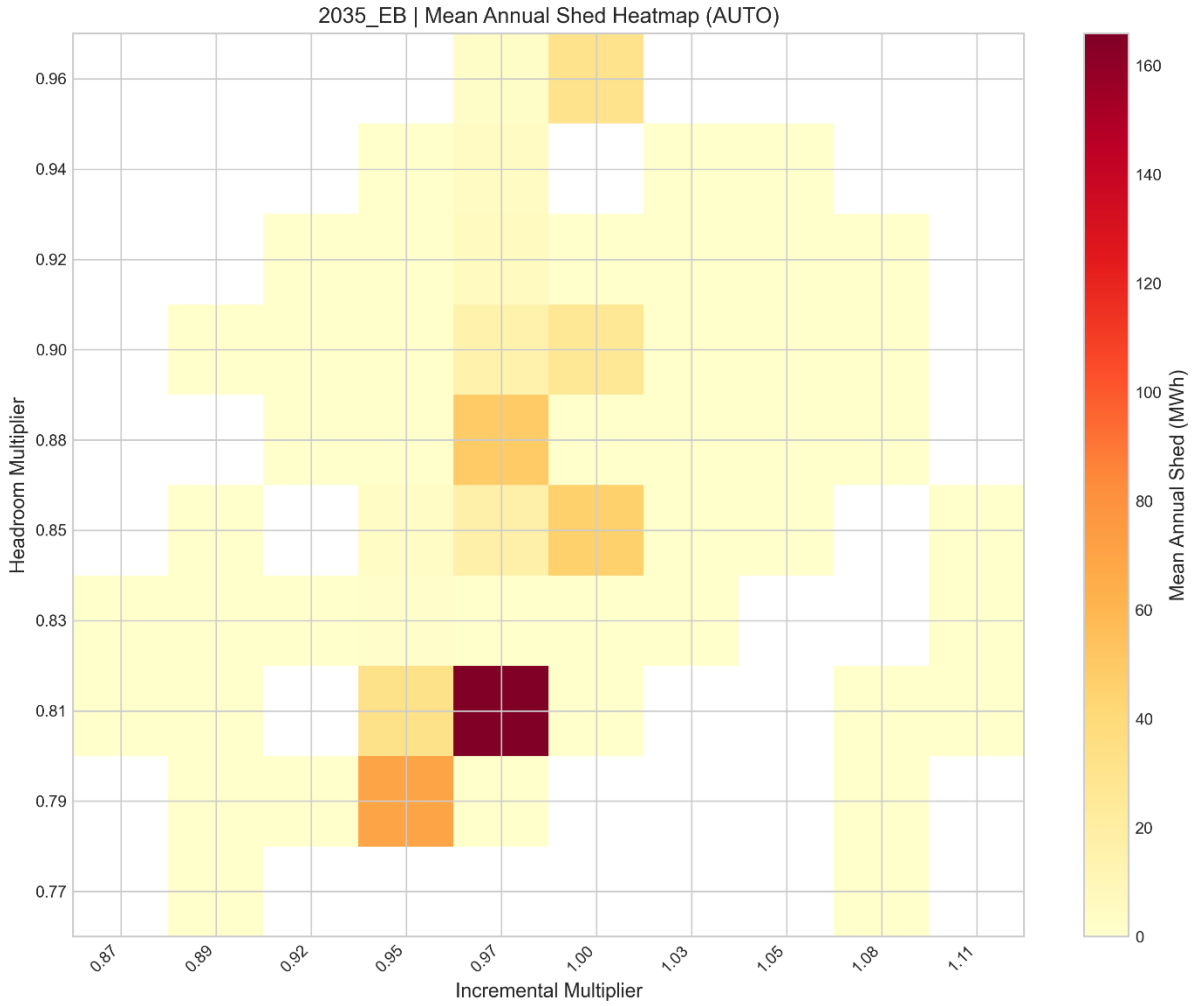


Figure 8.27 Mean annual shed under AUTO as a function of headroom and incremental-load multipliers (2035_EB)

For adaptive-planning interpretation, this is valuable because it translates the ensemble into observable signposts, where persistent movement toward lower headroom (leftward on the plot) and/or higher realised incremental load (rightward) increases the likelihood that the economically preferred choice shifts from S3 toward S4. Complementarily, if S3 were instead treated as a commit-

now fixed policy and evaluated under a practical tolerance (e.g., shed_fraction ≤ 0.001), then 34 of 100 futures are non-satisficing and cluster in a more adverse region (approximately U_headroom_mult ≈ 0.77–0.95 and U_inc_mult ≈ 1.02–1.13). Together, these diagnostics demonstrate how the screening layer probes policy vulnerability by identifying policy-relevant subsets of futures (vulnerability regions), rather than requiring exploration of the entire uncertainty space.

8.5.3 Decision-relevant outcomes

Across both pathways, the robustness screen produces decision-relevant conclusions that are not accessible from single-scenario feasibility checks.

Dominated strategies are eliminated decisively. For both 2035_BB and 2035_EB, low-uplift strategies (S0–S2) are consistently non-robust: they exhibit persistent unserved energy across the ensemble and therefore incur penalty-dominated grid-interface costs. This is an intended property of the screen. It functions as an early-stage filter that removes non-credible grid-interface policies before any attempt at tighter coupling or more detailed network representation.

Grid adequacy is pathway-conditional and becomes a policy choice under electrification. Under 2035_BB, S3 is robustly sufficient: it yields zero shedding across all 100 futures and is selected by AUTO in 100% of cases (Figure 8.22; Table 8.5), implying that further uplift (S4) is weakly dominated on cost grounds within the tested uncertainty set. Under 2035_EB, by contrast, adequacy is no longer invariant, where S3 and S4 represent a cost–reliability trade-off, and the least-cost choice switches across futures (Figure 8.24). This distinction is central for decision support and shows the electrification pathway can create conditions under which higher grid investment becomes economically justified, even if it is not uniformly required.

Scenario-discovery diagnostics expose vulnerability regions and support signpost design. The screening outputs do not merely report averages; they identify where policies become stressed. Figure 8.27 provides a compact “vulnerability map” that localises the subset of futures in which AUTO retains non-zero residual shortfall (and hence where the decision boundary lies between accepting small shortfall versus paying for additional uplift). In combination with the fixed-policy interpretation reported above (e.g., the S3 non-satisficing subset under a tolerance-based satisficing rule), the results illustrate a decision-analytic workflow: rather than attempting to span the entire uncertainty space, the model isolates a smaller, policy-relevant set of futures (vulnerability regions) that can be communicated transparently to stakeholders and operationalised as signposts (e.g., sustained headroom erosion and realised incremental-load growth).

Methodological capability is demonstrated with a thin-waist interface. Crucially, these decision insights are obtained without multi-pass optimisation: the site module exports a single incremental electricity trace, and the grid robustness screen consumes it alongside an uncertainty ensemble and a discrete upgrade menu. This preserves reproducibility and modularity while still enabling ensemble-based policy screening, regret-based benchmarking (AUTO), and scenario-discovery style diagnostics, capabilities that can later be deepened (e.g., richer network models or iterative coupling) without changing the site model's internal structure.

8.6 Site decision robustness overlay

Section 8.5 demonstrated that grid adequacy is pathway-conditional. Under 2035_BB, a moderate reinforcement option (S3) is robustly sufficient across the tested futures and is selected by the AUTO benchmark in all cases, implying that further uplift is weakly dominated within the PoC uncertainty set. By contrast, under 2035_EB the adequacy decision becomes economically salient: the least-cost benchmark switches between S3 and S4 across futures (approximately 51% vs 49%), reflecting a trade-off between paying a higher annualised uplift charge and avoiding residual tail penalties in adverse headroom conditions.

This section therefore moves from pathway-conditional screening to an explicit site pathway comparison between 2035_EB and 2035_BB, evaluated over the same 100 paired futures. The analysis applies the site decision robustness overlay defined in Chapter 7, a deliberately lean post-processing layer that does not re-solve dispatch and does not introduce multi-pass coupling, but instead (i) re-costs a small set of annual site cost components using per-future multipliers and (ii) adds the pathway-consistent grid screening cost under the AUTO benchmark. Paired-futures comparability is enforced by joining results explicitly on *future_id*, ensuring that any EB–BB ranking differences arise from pathway structure and interface consequences rather than from different uncertainty draws.

The purpose of the overlay is thus decision-analytic rather than operational, to test whether one pathway robustly dominates the other in an ex-post accounting sense (win rates and regret distributions), and to identify the subset of futures in which the ranking flips, thereby revealing which uncertainties most strongly govern the economic preference between electrification and biomass completion under the PoC's one-pass architecture.

8.6.1 Overlay construction

The decision evaluated in this section is intentionally narrow and decision-facing: select the preferred 2035 site pathway, either 2035_EB (electrification completion) or 2035_BB (biomass completion). The comparison is conducted under a paired-futures contract, meaning both pathways are evaluated over an identical ensemble of 100 futures and compared *one-to-one* using the same *future_id*. This prevents pathway ranking being distorted by different uncertainty realisations and ensures that any difference in outcome is attributable to pathway structure and its interface consequences rather than sampling artefacts.

Pathways are ranked using a single objective metric, the future-scaled total system cost, constructed by combining:

- an annually summarised site cost basis taken from the deterministic dispatch artefacts for each pathway;
- an accounting overlay that applies per-future multipliers to selected site cost components (electricity, biomass, and an ETS proxy); and
- a grid-interface consequence taken from the grid robustness screen outputs, implemented as the per-future AUTO (min-cost) response within the candidate upgrade set (including annualised uplift cost and any residual reliability penalty captured by the screening proxy). The critical interpretive point is that the overlay is accounting-only, re-weighting annual cost components without re-optimising dispatch and without altering operational artefacts.

To make this framing auditable, the paired comparison is recorded directly in *site_decision_robustness_2035_EB_vs_2035_BB.csv*, which joins EB and BB results by *future_id* and reports the resulting system-cost ranking and regret fields. The uncertainty multipliers used in the overlay (and in the grid screen) are defined in *futures.csv*. Table 8.7 summarises the fixed pathway artefacts from deterministic dispatch summaries that are held constant across all futures.

Table 8-7 Fixed pathway artefacts re-costed by the overlay

Quantity used in overlay	2035_EB	2035_BB	Interpretation in the overlay
Baseline annual site cost (NZD m)	91.00	76.83	Annual site cost basis before future multipliers (from dispatch totals/KPIs)
Annual electricity use (GWh)	827.27	592.94	Physical electricity quantity used to scale electricity-cost component
Site CO_2 (tCO_2/yr)	0.00	1,610.28	Physical emissions quantity used by the ETS proxy component

Table 8.8 reports, the multiplier ranges actually explored in the 100 paired futures (values shown as P10, Median, and P90 with Min–Max).

Table 8-8 Uncertainty multipliers used in the paired-futures ensemble

Multiplier	Sensitivity Analysis	P10 / Median / P90 (Min–Max)
Electricity price multiplier P_{elec}	Site electricity cost component	0.785 / 0.987 / 1.184 (0.687–1.293)
Biomass price multiplier P_{bio}	Site biomass cost component	1.133 / 1.486 / 2.038 (0.958–2.327)
ETS proxy multiplier P_{ETS}	ETS/carbon proxy cost component	0.950 / 1.275 / 1.654 (0.835–1.891)
Headroom multiplier U_{head}	Grid screen: effective headroom (adequacy stress)	0.817 / 0.890 / 0.950 (0.765–0.987)
Incremental-load multiplier U_{inc}	Grid screen: scaling of site incremental load	0.927 / 1.003 / 1.081 (0.867–1.133)
Upgrade CAPEX multiplier $U_{upgrade}$	Grid screen: annualised upgrade cost	0.920 / 1.071 / 1.285 (0.830–1.453)
VOLL U_{VOLL}(NZD/MWh)	Grid screen: penalty on residual shedding	10,000 / 15,000 / 20,000 (5,000–20,000)
Consenting uplift $U_{consent}$	Grid screen: consenting/cost uplift factor	1.125 / 1.252 / 1.422 (1.028–1.561)

Together, these artefacts ensure the comparison isolates how economic and planning assumptions alter preference ordering across the same operational evidence, rather than implying that the overlay changes feasibility or dispatch behaviour.

8.6.2 Ensemble comparison

This subsection presents the core comparative result of the site decision robustness overlay, and evaluates across the same 100 paired futures, which 2035 pathway is preferred, and how strongly? Preference is assessed using the future-scaled total system cost (site accounting overlay + grid-interface consequence under the AUTO benchmark), and robustness is assessed through dominance frequency (win rate) and regret, interpreted as the penalty incurred if a pathway is chosen ex ante and the realised future favours the alternative.

Figure 8.28 Paired-futures system cost distribution (2035_EB vs 2035_BB) compares the distribution of total system cost across the paired futures. Two findings are immediate. First, 2035_BB is consistently cheaper in the central mass of the ensemble. The median system cost is NZD 86.83 m for BB versus NZD 96.26 m for EB (Table 8.9), corresponding to a median paired-future gap of approximately NZD 9.28 m (EB more expensive). This separation is substantial for a PoC-scale screening layer and indicates that the pathway ordering is not driven by small perturbations but reflects a systematic difference in how the overlay re-costs the two pathways under identical futures.

Second, 2035_EB exhibits a heavier upper tail, consistent with its greater exposure to (i) electricity-related overlay multipliers (given higher electricity consumption under EB) and (ii) pathway-conditional grid-interface costs (AUTO switching between S3 and S4 for EB, versus stable S3 for BB). In the upper decile, EB reaches P90 = NZD 113.94 m, compared with P90 = NZD 101.06 m for BB (Table 8.9). Thus, EB is not only typically more expensive; it also carries a materially larger adverse-cost tail within the tested uncertainty ranges.

A decomposition of the median paired-future gap confirms that the overlay acts primarily as an economic ranking device rather than a second feasibility test. The median EB–BB difference is dominated by the site-side accounting overlay (≈ +7.96 m NZD), with an additional contribution from the grid-interface consequence (≈ +1.24 m NZD) arising from the AUTO-selected adequacy response embedded in the robustness screen.

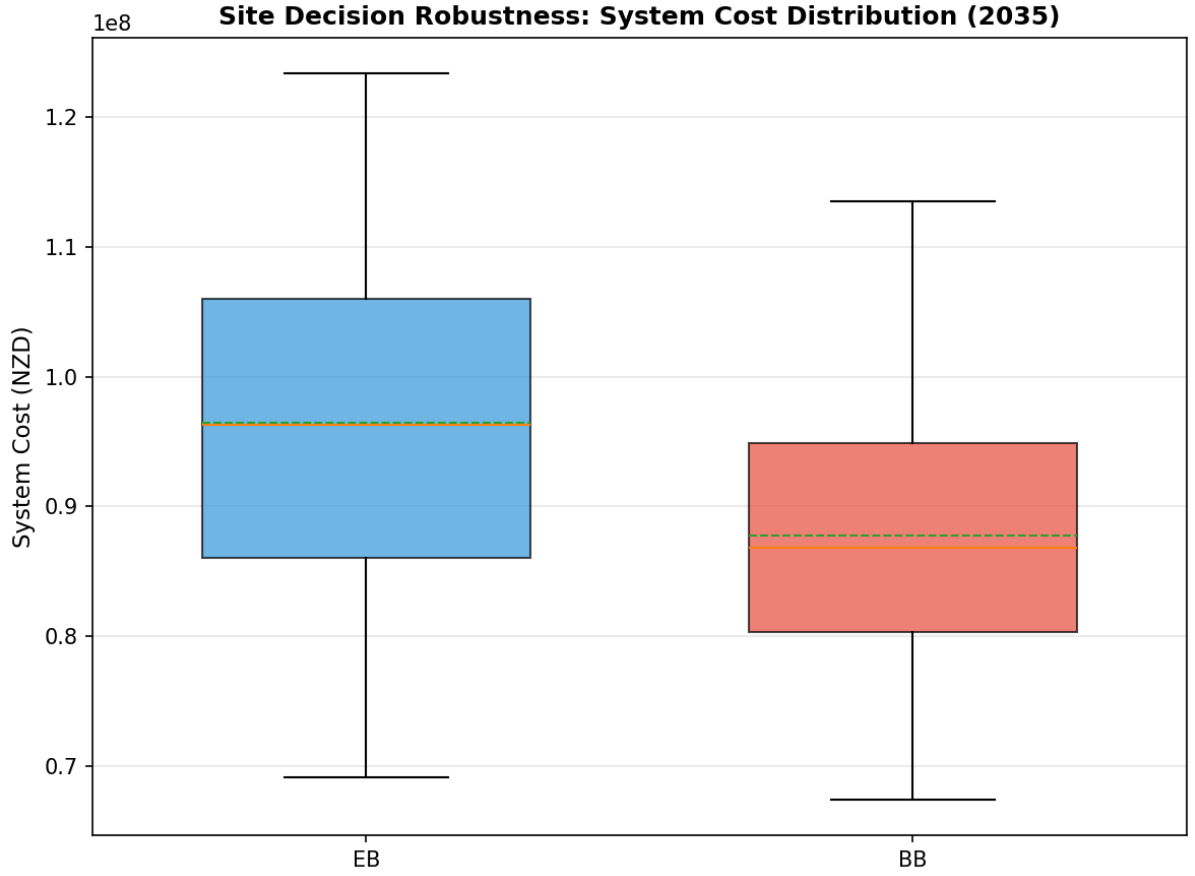


Figure 8.28 Paired-futures system cost distribution (2035_EB vs 2035_BB); values in NZD ($\times 10^8$); boxes show the interquartile range, whiskers the full range

The two components behave differently. The site-side component reflects relative fuel and electricity prices and is present in every future. The grid-interface component is threshold-triggered: under

2035_BB the AUTO response is S3 in every future, whereas under 2035_EB it moves to S4 in 49 of the 100 futures (Section 8.5.2), and the grid-interface difference between the pathways arises mainly in those futures. Extending the boundary to the grid interface therefore did not change the ranking in this case, but it made visible a consequence that a site-boundary comparison cannot represent whether, and how often, the electrification choice triggers the next reinforcement step.

Figure 8.29 Regret CDF under the site decision overlay (2035_EB vs 2035_BB) reframes the comparison as a commit-now robustness test. Regret is interpreted as the penalty incurred if a pathway is selected ex ante and the realised future would have favoured the alternative. The result is sharply asymmetric. BB has zero regret in 93% of futures, because it is the minimum-cost pathway in those futures; its empirical CDF therefore rises almost vertically at (or extremely close to) zero. EB, by contrast, incurs positive regret in 93% of futures, with regret accumulating into a long tail. The median regret of choosing EB is NZD 9.28 m, with P90 = NZD 15.82 m and a worst-case regret = NZD 18.93 m (Table 8.9). BB is not "risk-free"—there are seven futures where EB is cheaper—but BB's downside from "being wrong" is comparatively limited, with worst-case regret = NZD 4.49 m.

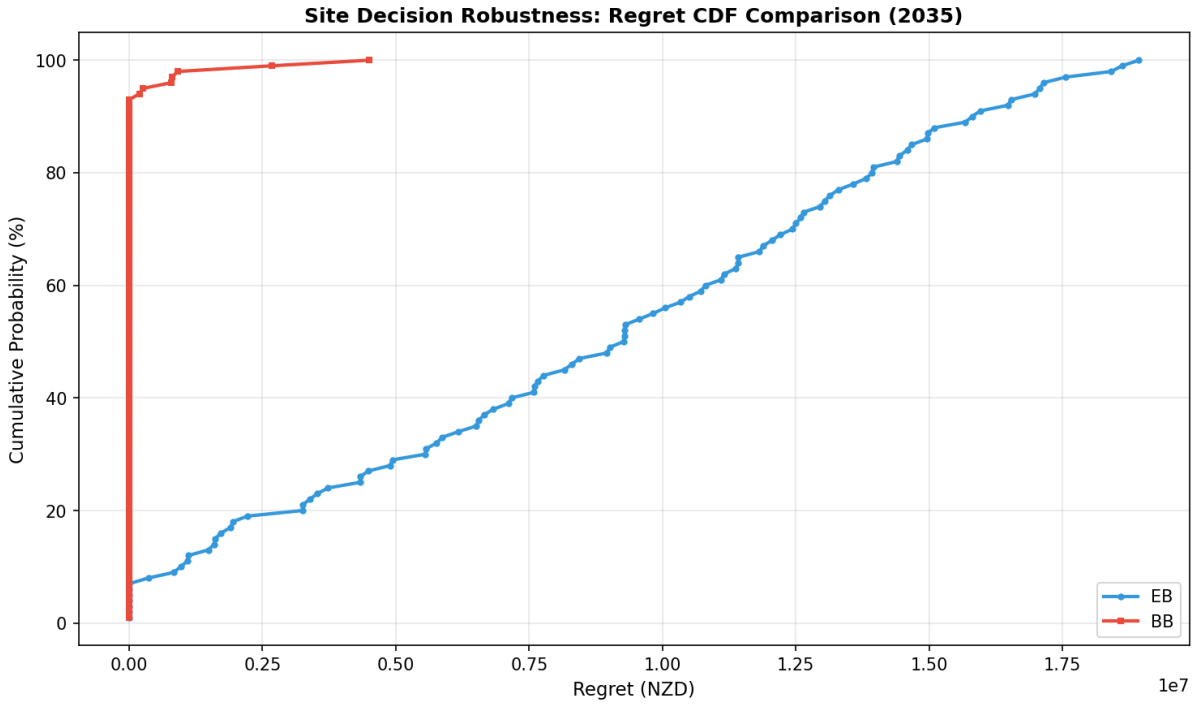


Figure 8.29 Regret CDF under the site decision overlay (2035_EB vs 2035_BB); regret in NZD ($\times 10^7$)

Table 8.9 consolidates these dominance and tail-risk statistics. The regret asymmetry provides the cleanest robustness interpretation: 2035_BB is not strictly dominant (EB wins 7% of futures), but BB is robustly preferred under the PoC overlay because it combines (i) a very high win rate, (ii) near-zero

regret in almost all futures, and (iii) a markedly lower worst-case regret. EB is therefore non-robust as a commit-now choice in this ensemble: selecting EB implies accepting a high probability of substantial regret relative to the best pathway under the realised future.

Table 8-9 Site decision robustness summary across 100 paired futures

Metric	2035_EB	2035_BB
Win rate (minimum-cost pathway share)	7%	93%
Median system cost (NZD m)	96.26	86.83
P90 system cost (NZD m)	113.94	101.06
Max system cost (m NZD)	114.62	97.36
Median regret if chosen ex ante (NZD m)	9.28	0.00
P90 regret if chosen ex ante (NZD m)	15.82	0.00
Max regret (NZD m)	18.93	4.49
Upgrade exposure to S4 (+150 MW) under AUTO	49%	0%
Satisficing rate under (S_{sys})	35%	88%

Taken together, Figures 8.28–8.29 and Table 8.9 demonstrate that, within the PoC's one-pass architecture and accounting-only uncertainty propagation, 2035_BB is robustly preferred over 2035_EB, not because EB never wins, but because EB's wins are rare and its regret is typically large when it does not win. The next subsection (8.5.3) characterises when the ranking flips and which uncertainties drive that reversal, converting the dominance statement into an interpretable "vulnerability map" for stakeholder discussion.

8.6.3 Flip futures and key drivers

While the ensemble comparison shows a strong overall preference for 2035_BB (93%-win rate), the paired-futures framing enables interrogation of the exceptions. In this ensemble, 2035_EB is cost-minimising in only seven futures (*future_id* ∈ {4, 21, 39, 40, 82, 83, 98}). These "flip" futures are not random anomalies; they display a coherent uncertainty signature that is informative for decision-making.

Table 8.10 provides a compact flip-region diagnostic. It reports min–max ranges for each uncertain driver separately for EB-win and BB-win futures, and an indicative salience ranking based on standardised coefficients from a simple linear diagnostic of the paired-future cost difference, ΔC^{sys}(EB − BB). Coefficients are sorted by absolute magnitude; positive values tend to penalise EB (increase ΔC^{sys}), while negative values tend to favour EB (decrease ΔC^{sys}).

Table 8-10 Flip-region diagnostic and uncertainty salience (futures where EB is cheaper vs BB is cheaper; sorted by impact on ΔC^{sys})

Uncertain driver	EB cheaper (n=7): min–max	BB cheaper (n=93): min–max	ΔC^{sys}
P_{bio} (biomass price multiplier)	1.995–2.258	0.958–2.327	−0.70
P_{elec} (electricity price multiplier)	0.729–0.978	0.687–1.293	+0.66
U_{inc} (incremental-load multiplier)	0.895–1.042	0.867–1.133	+0.14
U_{head} (headroom multiplier)	0.824–0.914	0.765–0.987	−0.04
U_{VOLL} (VOLL, NZD/MWh)	10,000–20,000	5,000–20,000	+0.02
P_{ETS} (ETS/CO_2 proxy multiplier)	0.950–1.733	0.835–1.891	−0.01
D_{heat} (heat-demand multiplier)	0.977–1.057	0.960–1.088	+0.01
U_{upgrade} (upgrade CAPEX multiplier)	1.000–1.255	0.830–1.453	+0.01
U_{consent} (consenting uplift)	1.132–1.369	1.028–1.561	+0.00

Two conclusions follow directly. First, the dominant drivers are site-side economics: the relative realisation of electricity and biomass price multipliers overwhelmingly governs the sign and magnitude of ΔC^{sys}. EB becomes competitive only in a "high-biomass/low-electricity" corner of the tested uncertainty space, consistent with the overlay structure (annual quantities are fixed; multipliers drive the economic ordering).

Second, grid-side multipliers are secondary but decision-relevant near the boundary. Among grid-related uncertainties, the incremental-load multiplier U_{inc} provides the clearest additional contribution to ΔC^{sys}, with headroom U_{head}playing a smaller role. This is consistent with Section 8.4: EB induces upgrade switching pressure (S3 ↔ S4) under adequacy uncertainty, and that pressure becomes consequential precisely when site-side economics make the two pathways close.

Figure 8.30 Winner map in price space (P_{elec} vs P_{bio}, coloured by preferred pathway) visualises this flip region. EB-win futures cluster where biomass is realised as expensive and electricity as cheap; outside this corner, BB remains preferred. This figure provides the most defensible PoC-scale diagnostic because it displays the flip pattern directly, without imposing a more complex classifier.

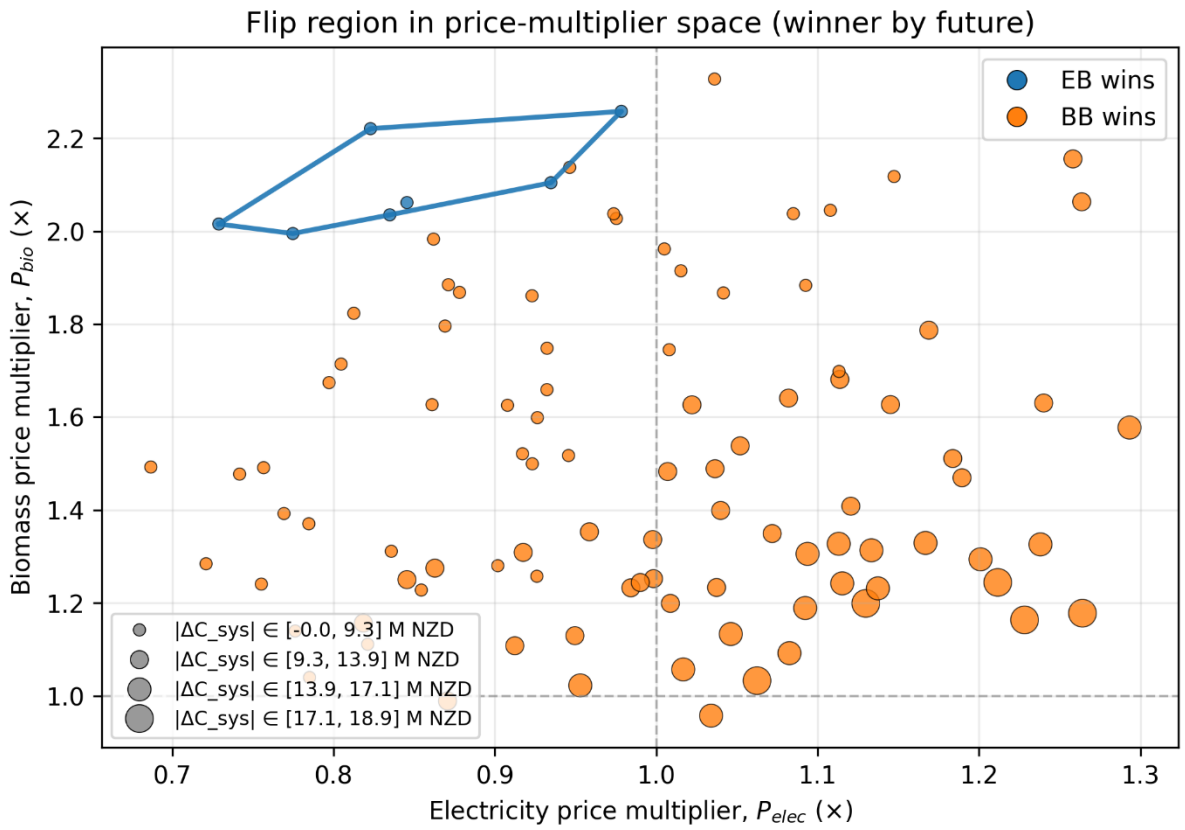


Figure 8.30 Winner map in price space (P_{elec} vs P_{bio}, coloured by preferred pathway)

The grid-interface component does not overturn the overall ordering, but it shapes the boundary of the flip region in a decision-relevant way. Two observations support this interpretation. (i) EB cannot "win on the grid" if it loses on the site: there are no futures in which the site overlay favours BB, yet EB becomes cheaper once grid-interface costs are added; favourable adequacy conditions are not sufficient to compensate for unfavourable site-side price realisations. (ii) Grid conditions can still penalise EB in near-boundary cases: there exist futures in which the site accounting component marginally favours EB, but BB remains cheaper in total system cost once the AUTO grid consequence is included. These cases occur near the electricity–biomass trade-off boundary and are characterised by incremental-load scaling that increases EB's adequacy cost exposure under AUTO (often pushing EB towards the higher-uplift option). This behaviour is consistent with Section 8.4's finding that EB is more prone to upgrade switching under headroom uncertainty.

Having established the paired-future comparison and the conditions under which the EB–BB ranking flips, the next section places these results in a broader interpretation of what the PoC demonstrates, and what it does not, particularly regarding the limits of accounting-only overlays, the role of benchmark grid responses, and how a richer model could endogenies some of the drivers identified here in future work.

8.7 Cross-cutting discussion

This section synthesises the PoC evidence against the thesis's core methodological objectives, to (i) develop an interpretable site to region modelling architecture, (ii) operationalise it as a governed computational environment, and (iii) demonstrate decision relevance under deep uncertainty consistent with DMDU/RDM principles (Chapters 3–5). The discussion follows a claims → evidence → limits structure. It highlights what the PoC artefacts now substantiate empirically and delineates the boundaries within which those claims remain valid.

8.7.1 Thin-waist modularity and interpretability

The PoC empirically validates the thin-waist design introduced in Chapter 6: modules interact through a deliberately minimal coupling contract (a small number of canonical signals), enabling site-level temporal detail to inform regional evaluation without collapsing either layer into an opaque joint solve. This design directly addresses a practical constraint in the New Zealand process-heat context: limited access to site-resolved operational and network data, and therefore heavy reliance on annualised (aggregated) evidence (e.g., RHDD summaries and RETA-type regional assessments).

Three aspects of Chapter 8 outputs support the claim:

1. **Stable deterministic artefacts across epochs.** DemandPack outputs preserve a consistent "shape language" across 2020–2035, supporting face-validity checks before introducing robustness claims (Figures 8.1–8.5). This stability allows the thesis to treat hourly structure as a disciplined extension of annual evidence, rather than an ad hoc construction.
2. **Decomposable dispatch evidence.** The dispatch layer remains decision-facing and interpretable; unit-resolved contributions (Figures 8.6–8.10), duty-cycle behaviour (Figures 8.11–8.14), and KPI summaries (Table 8.2) provide mechanistic explanations for pathway differences rather than relying on single aggregate metrics. Table 8.1 (run ledger) anchors these outputs to specific run identifiers (e.g., *poc_20260105_release02*, *dispatch_prop_v2_capfix1*), strengthening auditability.
3. **Minimal coupling interface with a single exported signal.** The regional screen consumes the exported incremental electricity trace, Δelectricity (Figures 8.15–8.16; Table 8.3), evaluates it against headroom evidence (Figures 8.17–8.18), and propagates consequences through the robustness screen (Section 8.4). This isolates interdependence in an explicit numerical object rather than embedding it implicitly in a monolithic model.

The PoC demonstrates interpretability under a constrained contract: one-way coupling (site → grid), a screening proxy rather than network physics, and proportional dispatch rather than scheduling. These

constraints define the maturity level of the implementation, not the viability of the architectural principle.

This modularity is only useful if results remain repeatable as assumptions and inputs evolve; the next subsection shows how the PoC treats reproducibility as a governed property of the environment rather than as a reporting afterthought.

8.7.2 Governed reproducibility

The PoC treats reproducibility as a governed property of the modelling environment, enforced through acceptance gates and canonical artefact pathways. This governance directly mitigates a common regional-analysis failure mode: fragmented transformations between annual evidence, hourly simulation outputs, and uncertainty summaries—particularly salient in settings dominated by RHDD/RETA-style aggregation.

The evidence is procedural and empirical.

1. **Acceptance-gated completeness is enforced.** Chapter 7 defines reporting gates and artefact completeness requirements (Table 7.1). Chapter 8 operationalises these requirements through the run ledger (Table 8.1), which records, for each reported result set, the executed layers, run identifiers, the futures file, and the evaluated decision options. This provides a reproducible map from textual claims to on-disk artefacts.

2. **Canonical artefacts function as objects of record.** The key results in Sections 8.4–8.5 derive from a small set of canonical artefacts rather than discretionary manual transformations. Table 8.11 lists these objects of record.

3. **Provenance persists across layers.** The deterministic foundations (Figures 8.1–8.16; Tables 8.2–8.3) are carried forward into the robustness screen (Figures 8.19–8.27; Tables 8.4–8.6) and then into the paired-futures overlay (Figures 8.28–8.29; Tables 8.7–8.10). The reader can trace late-stage metrics back to operational artefacts rather than being asked to accept uncited aggregation.

Table 8-11 Canonical artefacts underpinning Chapter 8 claims

Layer	Canonical artefact (object of record)	Role in governed reproducibility
Futures	*Input/rdm/futures_2035.csv*	Fixes the uncertainty ensemble and paired-futures contract
Robustness screen	strategy/future matrices and screened cost outputs (Section 8.4)	Regenerable basis for Figures 8.19–8.27 and Tables 8.4–8.6
Site decision overlay	site_decision_robustness_2035_EB_vs_2035_BB.csv	One-to-one paired comparison joined on future_id
Overlay summary	site_decision_robustness_summary_2035.csv	Regenerable basis for Tables 8.7–8.10

Governance guarantees reproducibility and auditability, not universal correctness. It reduces epistemic fragility (irreproducible results), but it does not eliminate model-form uncertainty (results that may change under higher-fidelity physics, scheduling, or calibration). The PoC therefore supports methodological claims about provenance and disciplined uncertainty propagation, while remaining explicit about the conditionality of substantive outcomes.

Conformity to the artefact contract is also not the same as adequacy for every downstream use. The exported ΔE series is computed once per pathway from the deterministic dispatch; the 100 futures vary it only through the incremental-load multiplier applied in the screen. Screen outputs indexed by future therefore carry future labels, but the underlying signal, including its peak and P95, is pathway-level. Every artefact passes its gates, yet a downstream analysis that read a future-indexed output as a future-specific load profile would over-interpret it. Each reuse of an artefact therefore needs a statement of what the artefact can support for that use, in addition to the gates it has passed.

8.7.3 Decision relevance under uncertainty

The PoC demonstrates decision relevance under uncertainty using RDM-consistent outputs, dominance frequency, regret distributions, tail exposure, and option switching, while respecting the one-pass coupling boundary. This aligns with the DMDU framing in Chapters 5, where robustness is assessed through performance across futures and vulnerability characterisation rather than reliance on a single expected-value optimum [9], [298].

The PoC answers two decision questions in canonical decision-analytic form.

1. **Grid-interface adequacy is pathway-conditional and triggers economically salient switching.** Pathway-specific Δelectricity (Figures 8.15–8.16; Table 8.3) evaluated against headroom evidence (Figures 8.17–8.18) feeds the robustness screen (Figures 8.22–8.27; Tables 8.5–8.6). The screen reveals a decision-relevant behavioural signature: under 2035_BB, the AUTO

benchmark (per-future least-cost option within the candidate upgrade set) selects S3 across the tested futures; under 2035_EB, AUTO switches between S3 and S4 across futures (≈51% vs 49%). This switching behaviour converts the abstract claim "electrification stresses the interface" into a traceable mapping from pathway → signal → upgrade-option sensitivity and tail penalty exposure.

2. **The 2035 pathway choice admits a paired-futures robustness ranking expressed in regret.** Section 8.5 compares 2035_EB versus 2035_BB under the same futures (future_id pairing), using system-cost distributions (Figure 8.28), regret CDFs (Figure 8.29), and a compact verdict table (Table 8.9 from site_decision_robustness_summary_2035.csv). The overlay further isolates the "flip futures" and identifies the dominant drivers of reversal (Section 8.5.3): primarily electricity versus biomass multipliers (site-side economics), with secondary influence from grid-side multipliers consistent with upgrade switching pressure.

The decision relevance extends beyond electricity, and the overlay explicitly includes biomass and ETS proxy terms. Even in an accounting-only treatment, these terms structure the conditions under which EB becomes competitive (high biomass scarcity/price and favourable electricity realisations) and provide a transparent handle for policy sensitivity, particularly relevant under evolving ETS settings and biomass supply uncertainty in regional transition narratives. The PoC does not claim to forecast these drivers; it demonstrates how the environment can test the decision implications of plausible volatility ranges in a disciplined way.

The decision relevance is bounded by three deliberate PoC constraints:

- One-pass coupling: the regional screen does not return scarcity or tariff signals to re-dispatch the site.
- Screening proxy: the grid layer acts as a guardrail diagnostic, not a planning-grade network model.
- Accounting overlay: multipliers re-cost fixed operational artefacts; they do not represent operational adaptation.

These bounds limit claims about coordinated equilibrium behaviour, but they strengthen interpretability and provenance, consistent with a post-positivist modelling stance where models act as structured instruments for learning under uncertainty rather than as final representations of reality.

8.7.4 Validity and PoC limitations

This subsection consolidates the main threats to validity so earlier results can be read as conditional, decision-relevant evidence rather than repeatedly caveated statements. The limitations are framed as opportunities for targeted refinement that preserve interfaces and governance.

1. **Synthetic demand fidelity (external validity).**

   Threat: stress-hour timing and peak coincidence may deviate from reality (Figures 8.1–8.5).

   Resolution path: targeted metering (representative seasonal weeks; import/steam logs) to calibrate timing/peaks without changing the artefact contract.

2. **Proportional dispatch (construct validity for operations).**

   Threat: does not represent scheduling, unit commitment, ramping, or operational adaptation (Figures 8.6–8.14; Table 8.2). Under proportional dispatch every electrode boiler operates whenever heat is demanded, so the 2035_EB export peaks at 133 MW (Table 8.3). The switching between S3 and S4 reported in Section 8.5.2 is therefore partly a property of the dispatch rule; a scheduled dispatch with a lower peak could reduce or remove the need for S4. This peak is well above the site import of about 85 MW indicated in the RETA Southland assessment [40].

   Resolution path: substitute a constrained LP/MILP or policy-informed dispatch kernel behind the same outputs (hourly unit contributions, KPIs, Δelectricity).

3. **Regional screening proxy (construct validity for network planning).**

   Threat: does not model network physics or operator planning processes (Section 8.4; Figures 8.19–8.27).

   Resolution path: progressively enrich the regional layer (e.g., PyPSA-style backend or simplified OPF) while preserving Δelectricity as the coupling signal.

4. **Interpretation of unserved/penalties (criterion validity).**

   Threat: unserved energy and penalties are conditional stress indicators, not outage forecasts.

   Resolution path**:** refine headroom evidence and constraint representation; keep unserved as an explicit stress metric with clear interpretation boundaries.

5. **One-pass coupling (internal validity for feedback dynamics).**

   Threat: cannot represent endogenous coordination between site operations and regional scarcity signals.

   Resolution path: introduce multi-pass iteration using the same signal pathway (return tariffs/shadow prices/caps), preserving the thin-waist contract and the governance already demonstrated.

6. **Boundary asymmetry (construct validity for the comparison).**

   Threat: the electrification pathway carries an explicitly modelled regional consequence (grid-interface headroom and reinforcement), whereas the regional consequences of the biomass pathway (availability, logistics, competing demand) enter only through a delivered-price multiplier (Section 7.4.2, boundary register). The comparison can show that the regional consequence of electrification is visible and evaluable; it cannot show that biomass completion is regionally unproblematic.

   Resolution path: a biomass supply module that returns a regional-consequence artefact (availability, delivered-cost curve, competing demand) through the same SignalsPack interface, so that both pathways are compared with regional consequences of equal standing.

Across these threats, the key methodological point is stable: improved data and higher-fidelity sub-models strengthen results without undermining the architecture, because interfaces and artefact pathways remain explicit and versionable.

8.7.5 Synthesis: what the PoC now justifies as the next step

The PoC substantiates the thesis's modular paradigm as an operational, governed approach to site to region decarbonisation analysis, and produces interpretable deterministic artefacts, propagates uncertainty through auditable screens and overlays, and yields decision-relevant robustness evidence early. This positions the framework as a practical complement to established national models (e.g., TIMES-NZ-style system optimisation), rather than competing on whole-system completeness, it adds site-to-region temporal granularity, explicit coupling signals, and RDM-style robustness screening, capabilities that are difficult to obtain from annualised regional narratives alone.

The PoC diagnostics do more than list possible extensions; they indicate an order. Four observations set it.

1. Biomass supply first. The flip-future diagnostic (Table 8.10) shows that the ranking turns mainly on the relative electricity and biomass prices, and the biomass price is the least evidenced input in the comparison, represented only by a broad multiplier (0.90–2.50; Table D.3). Combined with the boundary asymmetry (threat 6), this makes regional biomass supply the refinement most likely to change the comparison. Hypothesis: representing biomass availability and delivered-cost curves will narrow, and may shift, the region of futures in which biomass completion is preferred.

2. Scheduled dispatch second. The 133 MW peak under proportional dispatch drives the S3/S4 switching that dominates the grid-interface consequence of electrification (Sections 8.5.2 and 8.7.4). Hypothesis: constrained LP/MILP dispatch will lower the exported peak and the frequency of switching to S4, while leaving the annual electricity quantity, and hence the site-side cost difference, largely unchanged.

3. A larger ensemble with scenario discovery third. With only 7 of 100 futures favouring electrification, the conditions for reversal cannot be characterised reliably at this ensemble size. Hypothesis: a larger, space-filling ensemble analysed with scenario discovery (PRIM/CART) will describe the electrification-favourable region as a compact, interpretable set of conditions on the price drivers.

4. Artefact qualification before reuse. Because exported signals are pathway-level although screen outputs are indexed by future (Section 8.7.2), each reuse of an artefact should be qualified for its intended use before downstream analyses depend on it.

The diagnostics locate a specific bottleneck: under electrification, it is the critical peak in exported electricity, rather than annual energy use, that triggers the step from the 97 MW to the 150 MW reinforcement (Sections 8.5.2 and 8.7.4). This points to measures that reduce or shift the peak, such as thermal energy storage and high-temperature heat pumps (Section 2.2.2). The higher efficiency of heat pumps would lower both the electricity demand and the peak exported to the grid interface. The proof of concept was framed around the site's committed and candidate units and does not evaluate these options, so its two-pathway comparison should be read within that scope. As the framework matures, such technologies can enter the site module as additional units behind the same interfaces. That applies both at Edendale, where much of the demand is high-temperature steam, and at sites with lower-temperature demand that are well suited to heat pumps. Representing heat-pump performance through alternative COP formulations (structural uncertainty) and ranges of COP values (parametric uncertainty) would allow their regret and satisficing performance to be evaluated

alongside the grid-interface consequences they create, linking site decisions more closely to network and wider sectoral constraints.

The remaining extensions (demand calibration, a network-aware regional layer, multi-site coincidence and multi-pass coupling) stay on the roadmap (Section 9.5), but the PoC diagnostics do not yet indicate that they would change the comparison as much as the first three.

Chapter 9
Conclusions and Perspectives

Industrial process heat is among the most complex and least tractable domains of contemporary energy systems, characterised by diverse technologies, site-specific configurations, and tight dependencies on regional fuel supply, infrastructure, and policy settings. Unlike the electricity sector, where balancing and market rules are centralised, process heat transitions arise from distributed, path-dependent choices made by individual firms under uncertainty, with decisions at specific facilities propagating through regional electricity and biomass systems (and other energy carriers) whilst regional constraints feed back into the credibility and cost of site-level options. Decarbonising this sector therefore demands an analytical architecture that links local operational realism with system-level constraints and reveals how strategies perform as external conditions shift, all while remaining tractable, transparent, and decision-relevant. The approach adopted here rests on a simple premise: carry complexity where it creates decision value and abstract it where it would only obscure traceability. This thesis set out to address a persistent gap in the modelling and governance of industrial process-heat decarbonisation in New Zealand: decisions at the facility scale (technology choice, operational flexibility, and sequencing) must increasingly be made under deep uncertainty and under regional constraints (grid headroom, fuel supply, consenting and infrastructure). Yet, the dominant evidence base and many prevailing analytical practices remain annualised, siloed, and weakly linked across scales. In this setting, a key methodological challenge is not simply to "increase detail", but to design an analytical environment that (i) remains interpretable and auditable when granular data are sparse, (ii) can propagate uncertainty in decision-relevant forms (dominance, regret, tail risk, switching), and (iii) can evolve progressively, incorporating improved data and richer physics, without breaking comparability, provenance, or stakeholder trust.

The thesis pursued this aim through four objectives:

1. **Conceptual Development:** synthesise philosophical and methodological insights (including deep uncertainty and robustness traditions) into design principles for next-generation energy models.
2. **Framework Design:** design a modular, site-to-region architecture that integrates site demand and dispatch with regional constraints in a transparent environment.
3. **Uncertainty-Aware Analysis:** embed exploratory methods (stochasticity, ensembles, robustness screening) suitable for deep uncertainty.

4. **Application and Learning:** demonstrate the framework through a proof-of-concept (PoC) application and extract methodological and policy-relevant learning.

Chapters 2–6 establish the conceptual and design basis for a modular decision environment; Chapters 7–8 operationalise and evaluate that basis through the PoC and its results. This chapter consolidates what the thesis contributes, what the PoC does (and does not) establish empirically, and how the work provides a credible roadmap towards a "flexible and evolving" modelling environment that can mature from a bounded demonstration into a progressively expanding decision laboratory for New Zealand's industrial transition to low carbon choices.

9.1 Contributions of the thesis

The thesis reframes energy-system modelling for industrial process heat from a predominantly forecast-and-optimise paradigm toward an explicitly decision-centred paradigm, termed Decision-Centred Modelling (DCM). The conceptual contribution is to treat the modelling system as a governed environment that supports learning, stress-testing, and least-regret reasoning under deep uncertainty, rather than as a single monolithic representation of "the" future system. This aligns with the DMDU/RDM tradition, where robustness is assessed through performance across plausible futures and where interpretability, vulnerability characterisation, and adaptive sequencing are central outcomes (e.g., RDM; DAPP; scenario discovery). The thesis argues that this orientation is particularly appropriate for New Zealand's process-heat transition, given heterogeneous sites, uneven regional constraints, evolving policy settings, and limited access to operationally granular datasets. Its core rule is that the analytical boundary follows the decision: it extends to the consequences that can change the comparison of alternatives, at whatever scale they arise, and no further. The contributions that follow correspond to the remaining DCM commitments: thin-waist artefact coupling, robustness evaluation under deep uncertainty, and progressive, non-destructive refinement (Section 9.4).

The principal architectural contribution is the thin-waist, site-to-region modular design. Within this architecture, each layer can be improved independently, while interdependence is carried through a small number of canonical exchange artefacts (signals) rather than through tightly coupled joint solves (while retaining the capability to introduce tighter coupling where warranted by data and decision value). The PoC instantiates this architecture with a deliberately minimal coupling contract (notably Δelectricity exported from site to grid screening), while maintaining explicit artefact governance (run ledgers, canonical CSV outputs, and acceptance-gated reporting). This contribution is methodological and demonstrates a practicable way to connect facility-scale temporal detail to regional evaluation under data sparsity, without requiring a full digital twin or network-grade physics at the initial stage.

The thesis contributes a pragmatic uncertainty-handling strategy appropriate for early-stage, data-constrained modelling: uncertainty is propagated through (i) a robustness screen at the site–grid interface, and (ii) a paired-futures, post-processing overlay for pathway comparison. Importantly, this uncertainty propagation is disciplined and auditable and produces canonical robustness outputs (dominance frequency, regret distributions, tail exposure, option switching), while explicitly avoiding claims that would require endogenised feedback, network physics, or re-optimised operations. The PoC therefore demonstrates that meaningful, robust insights can be produced early, provided interpretation boundaries are explicit and provenance is governed.

The PoC provides an empirical demonstration anchored in a New Zealand industrial setting. It yields concrete learning about (i) how site pathway choices translate into time-resolved regional interface stress, (ii) how uncertainty shapes upgrade switching behaviour, and (iii) under what futures electrification completion (2035_EB) versus biomass completion (2035_BB) becomes economically preferred within the PoC accounting. Equally importantly, the PoC identifies where additional data and modelling fidelity would most reduce decision uncertainty, turning “limitations” into a targeted calibration and extension agenda.

9.2 Summary of key findings from the PoC

The PoC evidence supports three substantive findings, each bounded by the PoC’s one-pass and proxy assumptions. First, deterministic site artefacts are stable and interpretable across epochs. The DemandPack and proportional dispatch outputs remain interpretable and comparable across epochs (2020, 2025, 2028, 2035 variants), without interface drift. This matters because it establishes a credible deterministic “backbone” before robustness claims are made, and seasonal envelope, peakiness, and utilisation narratives can be inspected directly in the artefacts. In a context where much available evidence is annualised, the ability to generate stable hourly artefacts from sparse inputs, while keeping the transformation transparent, is itself a methodological achievement and a prerequisite for meaningful site-to-region coupling. The boundaries here are: the demand profiles are synthetic, and the dispatch kernel is proportional (rather than scheduling-grade); therefore, the PoC supports claims about interpretability and traceability of a deterministic core, not claims of fully calibrated operational realism.

Second, the robustness screen demonstrates that grid-interface adequacy is conditional on the exported site signal and uncertain headroom conditions. Under 2035_BB, the 97 MW reinforcement option is sufficient in every future within the PoC uncertainty set; under 2035_EB, the least-cost response switches to the 150 MW option in 49 of the 100 futures. This reflects an economically

meaningful trade-off between paying higher annualised grid upgrade costs and avoiding tail penalties under adverse headroom and incremental-load multipliers. This is decision-relevant and converts a qualitative statement ("electrification increases grid stress") into an auditable screening result about when additional reinforcement becomes justified and how often decision-makers would rationally switch between options under uncertainty. This is a guardrail screen, not a planning-grade network model; it quantifies conditional adequacy exposure given assumed headroom evidence, upgrade menus, and penalty proxies.

Third, the paired-futures overlay compares 2035_EB and 2035_BB over identical futures (same *future_id*), using future-scaled system cost that combines (i) site cost components re-costed by multipliers (electricity, biomass, ETS proxy) and (ii) the pathway-consistent grid-interface consequence under the AUTO benchmark. Within the PoC's uncertainty ranges, 2035_BB is preferred in 93 of the 100 futures and exhibits low regret as a "commit-now" choice, while 2035_EB wins only in a small subset of futures characterised primarily by favourable electricity versus biomass realisations, with grid-side multipliers exerting secondary influence consistent with the observed switching pressure. This finding is valuable not because it closes the pathway debate, but because it demonstrates that the environment can produce a reproducible vulnerability map and identifies which uncertainties actually govern preference reversal, and therefore which uncertainties deserve attention in stakeholder discussion, data collection, and policy sensitivity work. The overlay is accounting-only and does not re-optimise dispatch; therefore, it supports statements about economic ranking under fixed operational artefacts, not claims that operations would remain unchanged under alternative prices or constraints. Because the regional consequences of biomass supply enter only through price, the finding establishes that the regional consequence of electrification is visible and evaluable, not that biomass completion is regionally adequate.

9.3 Implications for energy transition analysis and governance in New Zealand

A central implication is that the proposed environment is designed to complement, not replace, existing national-scale models (including TIMES-type optimisation frameworks) and regional transition assessments (such as RETA-style work). Where national optimisation models provide system-level consistency and long-horizon investment insight, they often struggle to represent facility heterogeneity and operational temporal structure at scale. Where regional studies provide grounded narratives and constraints, they often lack a quantitative mechanism to translate site-specific pathway choices into time-resolved regional interface consequences under uncertainty.

The PoC shows a concrete bridge between these layers: a site pathway generates an explicit time-stamped carrier interaction (Δelectricity), which is then evaluated under uncertain headroom and policy menus. Extending this pattern beyond electricity would enable similarly explicit coupling for biomass constraints (regional availability caps, supply curves) and policy constraints (ETS-linked signals), turning currently siloed narratives into testable, auditable quantitative statements.

In data-constrained sectors, the practical question is not "what is the perfect model?", but "what data should be collected next to most improve decision confidence?". The thesis's environment provides a mechanism for answering that question systematically. By revealing which artefacts drive tail risk and preference reversal (e.g., peak timing in DemandPack; headroom multipliers near switching boundaries; electricity-versus-biomass multiplier relationship in flip futures), the environment turns limitations into a prioritised measurement and calibration plan. This is particularly important where datasets such as RHDD provide valuable coverage but limited operational granularity, therefore, the model can guide where incremental data effort yields maximum marginal value for robust planning.

Although the PoC deliberately avoids multi-pass coupling and endogenous adaptation, the demonstrated outputs align naturally with adaptive planning concepts: upgrade switching behaviour under uncertainty, regret distributions for commit-now choices, and explicit identification of vulnerability regions are the building blocks of signposts, triggers, and staged decisions. In other words, even a bounded one-pass implementation can support adaptive reasoning and can indicate where conservative "insurance" options become justified, and under what conditions a strategy should be reconsidered.

9.4 Limitations and validity: what remains conditional and why this is constructive

The thesis explicitly distinguishes governed reproducibility from empirical correctness: the PoC ensures the computational path is reproducible and auditable, but model-form uncertainty remains. The principal threats to validity are: (i) synthetic demand fidelity (stress-hour realism), (ii) proportional dispatch limits (scheduling and adaptation), (iii) screening proxy limits (network physics and operator planning processes), and (iv) one-pass coupling limits (absence of endogenous feedback), and (v) boundary asymmetry (the regional consequences of biomass supply are represented only through price). These limitations are decisive for interpretation, but they are not architectural defects; they are maturity-conditioned abstractions designed to be tightened progressively.

Crucially, the thin-waist design ensures that improvements, whether from targeted metering, refined dispatch constraints, richer network representations, or expanded ensembles, strengthen the outputs without breaking the logic, interfaces, or reporting governance. This "non-destructive refinement"

property is a key thesis contribution that enables the environment to evolve into a living modelling laboratory rather than a one-off case study.

9.5 Roadmap for progressive expansion: from PoC to an “alive” decision environment

Building directly on Chapters 5–8, the thesis supports a staged roadmap where each expansion passes through governed acceptance gates and preserves canonical artefact contracts. The near-term order follows the PoC diagnostics (Section 8.7.5): the refinements most likely to change the comparison come first.

Near-term extensions (next development cycle)

1. Regional biomass supply: represent biomass availability, delivered-cost curves and competing demand as a supply module returning a regional-consequence artefact through SignalsPack, removing the boundary asymmetry between the two pathways.
2. Dispatch refinement behind stable interfaces: replace proportional dispatch with constrained LP/MILP (unit constraints, ramping, minimum turndown, maintenance derates), while preserving outward-facing artefacts (unit heat supply traces, KPIs, Δelectricity), and test its effect on the exported peak and on upgrade switching.
3. Richer uncertainty ensembles and scenario discovery: expand futures with a space-filling design and apply scenario discovery (e.g., PRIM/CART) to derive interpretable vulnerability rules for the futures in which the preferred pathway reverses.
4. Artefact qualification for reuse: record, for each artefact, the downstream uses it can and cannot support, so that pathway-level signals are not read as future-specific evidence.
5. Targeted demand calibration: incorporate representative-week metering or operational logs to calibrate peak timing, seasonal envelopes and coincidence.
6. Multi-site scaling within a region: include additional industrial sites to quantify coincidence effects and shared reinforcement needs.

Medium-term extensions (regional realism and feedback)

1. Network-aware regional electricity layer: introduce a PyPSA-style or simplified OPF representation to replace or complement headroom proxies, while preserving Δelectricity as the thin-waist boundary variable.
2. Multi-pass coupling: return scarcity/tariff signals (time-of-use vectors, marginal adequacy prices, constrained-hour import caps) from the regional module to the site layer and iterate (Gauss–Seidel-style), enabling endogenous adaptation without architectural redesign.

3. Carrier expansion beyond electricity and biomass: couple process emissions constraints and ETS-linked signals in a way that preserves provenance and comparability.
4. Technology scope: add thermal storage and high-temperature heat pumps as site units, with COP treated as both a structural and a parametric uncertainty.

Long-term vision (national relevance and institutional adoption)

1. National-scale site aggregation and regional nesting: expand from single-site and single-region demonstrations toward a nested, multi-region implementation with consistent artefact contracts.
2. Institutional use as a validation gate: position the environment as a validation and stress-testing layer for broader planning work, enabling agencies to test how aggregated assumptions translate into operational stress and infrastructure exposure.
3. Extensibility across sectors: retain the thin-waist approach to extend beyond process heat (e.g., transport, buildings) while maintaining reproducibility and decision relevance under uncertainty.

9.6 Concluding remarks

This study advances a pragmatic but rigorous answer to a pressing methodological need in industrial decarbonisation planning: how to connect heterogeneous, high-resolution site behaviour to regional constraints and policy uncertainty in a way that remains transparent, auditable, and progressively improvable. The PoC does not claim final optimal pathways or planning-grade network conclusions; instead, it demonstrates that a governed, thin-waist modular environment can already produce decision-relevant robustness evidence (dominance, regret, tail exposure, and switching) despite severe data constraints.

By making coupling explicit through canonical signals and by treating reproducibility as a governed property of the modelling environment, the thesis establishes a foundation for an “alive” analytical platform: one that can mature through targeted calibration, expanded ensembles, richer physics, and multi-pass feedback without breaking comparability or trust, and that can accommodate emerging technologies and revised evidence while preserving auditability. In the longer term, the environment could also serve as a structured repository of modelled artefacts and decision traces, supporting cumulative learning across studies and facilitating extension to broader socio-economic dimensions where appropriate. In the New Zealand process-heat transition, where annualised evidence and siloed regional narratives remain prevalent, this constitutes a substantive methodological contribution: it

turns cross-scale decarbonisation analysis into a disciplined, cumulative learning process capable of producing increasingly confident and actionable insights as the environment evolves.

References

[1] MBIE and EECA, “Process Heat – Overview (Fact sheet),” 2018, [Online]. Available: https://www.mbie.govt.nz/dmsdocument/152-process-heat-current-state-fact-sheet-pdf

[2] I. Internationl Energy Agency, “Net Zero Roadmap,” 2023.

[3] IEA, “Global Energy Review 2025,” 2025. [Online]. Available: www.iea.org

[4] Climate Change Commission, “Advice on NZ ETS unit limits and price control settings for 2025-2029,” no. February 2024, 2024, [Online]. Available: https://www.climatecommission.govt.nz/our-work/advice-to-government-topic/nz-ets/our-advice-on-the-nz-ets/nzets-advice-2025-29/

[5] MfE, *Aotearoa New Zealand’s first emissions reduction plan. Table of actions.* 2022.

[6] MBIE, “Process Heat in New Zealand: Opportunities and barriers to lowering emissions,” 2019. [Online]. Available: www.mbie.govt.nz.

[7] EECA, “Regional Heat Demand Database,” no. November 2022, pp. 1–10, 2024, [Online]. Available: https://www.eeca.govt.nz/insights/data-tools/regional-heat-demand-database/

[8] Enviser Limited and Emission Impossible Limited, “Annual Greenhouse Gas Emissions in Canterbury 2018 and 2021,” 2023.

[9] V. A. W. J. Marchau, W. E. Walker, P. J. T. M. Bloemen, and S. W. Popper Editors, *Decision Making under Deep Uncertainty*. 2019.

[10] R. J. Lempert, “Shaping the next one hundred years: New methods for quantitative, long-term policy analysis,” 2003.

[11] A. Herbst, F. Toro, F. Reitze, and E. Jochem, “Introduction to Energy Systems Modelling,” 2012.

[12] Energy Dep. US, “Industrial Process Heating -Technology Assessment,” 2015.

[13] J. A. Mathews, “Global trade and promotion of cleantech industry: a post-Paris agenda,” *Climate Policy*, vol. 17, no. 1, pp. 102–110, Jan. 2017, doi: 10.1080/14693062.2016.1215286.

[14] A. M. Lahijani, M. D. Protheroe, and M. Gschwendtner, “A modelling tool selection for decarbonising industrial process heat systems,” *Renewable and Sustainable Energy Reviews*, vol. 210, no. February 2024, p. 115149, 2025, doi: 10.1016/j.rser.2024.115149.

[15] EECA, “GIDI: Industrial Applicant Guide,” 2023.

[16] Energy Efficiency & Conservation Authority (EECA), “Regional Energy Transition Accelerator,” 2023, Accessed: Feb. 09, 2024. [Online]. Available: https://www.eeca.govt.nz/co-funding-and-support/products/about-reta/

[17] MBIE, “Energy in New Zealand 2024,” 2024.

[18] R. J. Lempert, *Decision Making under Deep Uncertainty From Theory to Practice*. 2019.

[19] J. H. Kwakkel, M. Haasnoot, and W. E. Walker, “Comparing Robust Decision-Making and Dynamic Adaptive Policy Pathways for model-based decision support under deep uncertainty,”

Environmental Modelling and Software, vol. 86, pp. 168–183, Dec. 2016, doi: 10.1016/j.envsoft.2016.09.017.

[20] S. S. Joe DeCarolis, “Project BlueSky: Building EIA’s Next Generation Energy Model,” 2024.

[21] S. Pfenninger, J. DeCarolis, L. Hirth, S. Quoilin, and I. Staffell, “The importance of open data and software: Is energy research lagging behind?,” *Energy Policy*, vol. 101, pp. 211–215, 2017.

[22] B. Fais, N. Sabio, and N. Strachan, “The critical role of the industrial sector in reaching long-term emission reduction, energy efficiency and renewable targets,” *Appl. Energy*, vol. 162, pp. 699–712, 2016, doi: 10.1016/j.apenergy.2015.10.112.

[23] R. Serth and T. Lestina, *Process heat transfer: Principles, applications and rules of thumb*. 2014. Accessed: May 02, 2023. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=phlk1tMcl9sC&oi=fnd&pg=PP1&dq=process+heat&ots=97OmGQxe-P&sig=0wyNZOMgc8InBnlCcQY4FtIns6M

[24] H. W. J. Rittel and M. M. Webber, “Dilemmas in a general theory of planning,” *Policy Sci.*, vol. 4, no. 2, pp. 155–169, 1973.

[25] I. E. Agency, “CO 2 Emissions in 2023,” *International Energy Agency*, vol. 24, p. 22, 2023, [Online]. Available: www.iea.org

[26] Energy Dep. US, “Innovating Clean Energy Technologies in Advanced Manufacturing Technology Assessments Process Heating Process Intensification Roll-to-Roll Processing Sustainable Manufacturing-Flow of Materials through Industry Waste Heat Recovery Systems Wide Bandgap Semiconductors for,” 2015.

[27] United Nations Climate Change, “nationally determined contributions (2024 NDC Synthesis Report).” [Online]. Available: https://unfccc.int/process-and-meetings/the-paris-agreement/nationally-determined-contributions-ndcs/2024-ndc-synthesis-report

[28] MBIE, “NEW ZEALAND ENERGY STRATEGY Context for an energy strategy,” 2022.

[29] EECA, “Decarbonising industrial process heat,” 2019.

[30] M. Atkins, “Options to Reduce New Zealand’s Process Heat Emissions,” 2019.

[31] I. Dincer and M. A. Rosen, *Exergy: energy, environment and sustainable development*. Elsevier, 2020.

[32] C. Arpagaus, F. Bless, M. Uhlmann, J. Schiffmann, and S. S. Bertsch, “High temperature heat pumps: Market overview, state of the art, research status, refrigerants, and application potentials,” *Energy*, vol. 152, pp. 985–1010, 2018.

[33] S. Madeddu *et al.*, “The CO2 reduction potential for the European industry via direct electrification of heat supply (power-to-heat),” *Environmental Research Letters*, vol. 15, no. 12, p. 124004, 2020.

[34] S. Lupi, “Fundamentals of electroheat,” *Electrical Technologies for Process Heating. Springer. Switzerland*, 2017.

[35] EECA, “High-temperature heat pumps for low carbon process heating,” 2019.

[36] IRENA, *Decarbonising hard-to-abate sectors with renewables*. 2024. [Online]. Available: https://www.irena.org/Publications/2024/Apr/Decarbonising-hard-to-abate-sectors-with-

renewables-Perspectives-for-the-G7#:~:text=Sectors that are particularly hard,fifth of total CO2 emissions.

[37] Energy Efficiency & Conservation Authority (EECA), “May 2024 Government Leadership Regional Energy Transition Accelerator (RETA) Bay of Plenty-Phase One Report,” no. May, 2024.

[38] Ministry for the Environment, “Our journey towards net zero,” 2024.

[39] EECA, “Accelerating the decarbonisation of Process Heat,” 2021.

[40] IPCC, “Technical Summary,” in *Climate Change 2021 – The Physical Science Basis*, Cambridge University Press, 2023, pp. 35–144. doi: 10.1017/9781009157896.002.

[41] Fonterra, “Fonterra ’ s Decarbonisation Journey,” 2024.

[42] EERE, “Improving Steam System Performance : A Sourcebook for Industry Second Edition,” 2012. [Online]. Available: https://energy.gov/sites/prod/files/2014/05/f15/steamsourcebook.pdf

[43] EECA, “Regional Energy Transition Accelerator (RETA) - Waikato – Phase One Report,” no. March, 2025, [Online]. Available: https://www.eeca.govt.nz/assets/EECA-Resources/Co-funding/RETA-Waikato-Phase-One-Report.pdf

[44] EECA, “Southland-Phase One Report October 2022 Government Leadership Regional Energy Transition Accelerator (RETA) Southland-Phase One Report.”

[45] Ernst and Young, “Regulatory options for New Zealand’s gas transmission system,” 2022.

[46] S. (IEA B. Peter Hall, “Residual biomass fuel projections for New Zealand ; 2024 Indicative availability by region and source October 2024 Residual biomass fuel projections for New Zealand ; 2024 Indicative availability by region and source,” no. October, 2024.

[47] Scion, “New Zealand Biofuels Roadmap Technical Report,” 2018.

[48] Beca, “Southland Murihiku Regional Energy Strategy 2022–2050. Summary.,” no. July, 2023.

[49] R. J. Lempert, “Decision Making under Deep Uncertainty and the Great Acceleration: The Role of Experts and Policy Analysis in a World in Transition,” *Decision Making under Deep Uncertainty and the Great Acceleration: The Role of Experts and Policy Analysis in a World in Transition*, no. April, 2025, doi: 10.7249/pea3789-1.

[50] M. Haasnoot, J. H. Kwakkel, W. E. Walker, and J. ter Maat, “Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world,” *Global Environmental Change*, vol. 23, no. 2, pp. 485–498, 2013, doi: 10.1016/j.gloenvcha.2012.12.006.

[51] J. R. Kasprzyk, S. Nataraj, P. M. Reed, and R. J. Lempert, “Many objective robust decision making for complex environmental systems undergoing change,” *Environmental Modelling and Software*, vol. 42, pp. 55–71, 2013, doi: 10.1016/j.envsoft.2012.12.007.

[52] M. for the Environment, “NES and NPS on industrial greenhouse gas emissions,” 2022.

[53] New Zealand. Ministry for the Environment, “New Zealand’s Initial Report under the Kyoto Protocol,” 2006.

[54] Energy and Efficiency Conservation Authority, Ministry for the Environment New Zealand, “National Energy Efficiency and Conservation Strategy,” no. September, pp. 4, 10, 2001.

[55] New Zealand Government, “Climate Change Response Act 2002,” no. 18 November 2002, pp. 1–319, 2002.

[56] P. Hodgson, “Renewable energy target for New Zealand,” no. April, 2002, [Online]. Available: http://beehive.govt.nz/release/renewable-energy-target-new-zealand

[57] H. Plant, S. Greater, T. One, and H. Kilowatts, “Heat Plant in New Zealand,” no. March, 2008.

[58] CRL Energy Ltd, “HEAT PLANT in NEW ZEALAND,” 2011.

[59] East Harbour Management Services and SCION, “Assessment of Possible Renewable Energy Targets,” 2007.

[60] C. Leining and S. Kerr, “Lessons Learned from the New Zealand Emissions Trading Scheme,” 2016. [Online]. Available: www.motu.org.nz

[61] MfE, *Phasing out fossil fuels in process heat : consultation document : national direction on industrial greenhouse gas emissions.* 2021.

[62] C. Leining, “A Guide to the New Zealand Emissions Trading Scheme: 2022 Update,” 2022. [Online]. Available: www.Motu.nz.

[63] C. Leining, S. Kerr, and B. Bruce-Brand, “The New Zealand Emissions Trading Scheme: critical review and future outlook for three design innovations,” *Climate Policy*, vol. 20, no. 2, pp. 246–264, 2020.

[64] R. Sims *et al.*, “Transition to a Low-Carbon Economy for New Zealand,” no. April, pp. 1–200, 2016.

[65] Transpower, “Taking the climate heat out of process heat,” 2019.

[66] C. A. Tracker, “New Zealand: Policies and Action.” Accessed: Jul. 22, 2025. [Online]. Available: https://climateactiontracker.org/countries/new-zealand/policies-action/

[67] Fonterra, “Accelerating renewable energy and energy efficiency,” 2020.

[68] Interim Climate Change Committee, “Action on agricultural emissions Accelerated electrification,” no. April 2019, 2019, [Online]. Available: www.iccc.mfe.govt.nz

[69] Concept Consulting, “Modelling the transition to a lower net emissions New Zealand,” 2018.

[70] M. Walmsley, T. Walmsley, L. Matthews, M. Atkins, J. Neale, and P. Kamp, “Pinch analysis techniques for carbon emissions reduction in the New Zealand industrial process heat sector,” *Chem. Eng. Trans.*, vol. 45, pp. 1087–1092, Oct. 2015, doi: 10.3303/CET1545182.

[71] B. N. (Energy Council), “Accelerating renewable energy and energy efficiency,” 2020.

[72] C. Kiro, “Resource Management (National Environmental Standards for Greenhouse Gas Emissions from Industrial Process Heat) Regulations 2023,” pp. 1–14, 2023.

[73] H. D. M. Woods, “Government ban on new coal boilers in place.” [Online]. Available: https://www.beehive.govt.nz/release/government-ban-new-coal-boilers-place#:~:text=The national direction will phase,from 27 July this year

[74] BlueScope, “Supporting New Zealand’s climate transition.” [Online]. Available: https://www.bluescope.com/our-steel/case-studies/supporting-new-zealands-climate-transition#:~:text=In 2023%2C we announced that,by the end of 2025.

[75] EECA, “Regional Energy Transition Accelerator He kupu whakataki,” no. June, 2025.

[76] Climate Change Commission, “Technical-Annex-Modelling-and-analysis-9-4,” 2024.

[77] Ministry for the Environment, “Annual updates to New Zealand Emissions Trading Scheme limits and price control settings for units 2025,” 2025. [Online]. Available: http://dx.doi.org/10.1016/j.bpj.2015.06.056%0Ahttps://academic.oup.com/bioinformatics/article-abstract/34/13/2201/4852827%0Ainternal-pdf://semisupervised-3254828305/semisupervised.ppt%0Ahttp://dx.doi.org/10.1016/j.str.2013.02.005%0Ahttp://dx.doi.org/10.10

[78] Anderson Lloyd, “Climate Change Response (Emissions Trading Scheme Agricultural Obligations) Amendment Act 2024,” 2024. [Online]. Available: www.al.nz

[79] M. De Witt, “Discretion in Decision Making: The Fonterra Case,” 2017.

[80] Changing Markets Foundation, “The New Merchants of Doubt Big Meat and Dairy’s Corporate Playbook to Distract, Delay, and Derail Climate Action,” 2024.

[81] V. Vijay, S. L. Pimm, C. N. Jenkins, and S. J. Smith, “The impacts of oil palm on recent deforestation and biodiversity loss,” *PLoS One*, vol. 11, no. 7, Jul. 2016, doi: 10.1371/journal.pone.0159668.

[82] Robin Martin (Taranaki–Whanganui regional correspondent), “Protesters call for Fonterra to phase out use of palm kernel expeller.” Accessed: Aug. 22, 2025. [Online]. Available: https://www.rnz.co.nz/news/country/533765/protesters-call-for-fonterra-to-phase-out-use-of-palm-kernel-expeller

[83] I. and E. (MBIE) Ministry of Business, “2324-1197 Reprioritising the Government Investment in Decarbonising Industry Fund,” 2023.

[84] Eloise Gibson, “Has cutting ‘corporate welfare’ left a hole in government climate plans?” Accessed: Aug. 23, 2025. [Online]. Available: https://www.rnz.co.nz/news/national/517269/has-cutting-corporate-welfare-left-a-hole-in-government-climate-plans

[85] New Zealand Productivity Commission, “Low-emissions economy,” 2018. [Online]. Available: https://www.productivity.govt.nz/sites/default/files/Productivity Commission_Low-emissions economy_Final Report_FINAL.pdf

[86] K. Levin, B. Cashore, S. Bernstein, and G. Auld, “Overcoming the tragedy of super wicked problems: constraining our future selves to ameliorate global climate change,” *Policy Sci.*, vol. 45, no. 2, pp. 123–152, 2012.

[87] C. Pickard and R. Pasqualino, “Long-Term Strategies for the Compatibility of the Aviation Industry with Climate Targets: An Industrial Survey and Agenda for Systems Thinkers,” *Systems*, vol. 10, no. 4, Aug. 2022, doi: 10.3390/systems10040090.

[88] T. Wiedmann, M. Lenzen, L. T. Keyßer, and J. K. Steinberger, “Scientists’ warning on affluence,” *Nat. Commun.*, vol. 11, no. 1, Dec. 2020, doi: 10.1038/s41467-020-16941-y.

[89] C. J. A. M. Termeer, A. Dewulf, and G. R. Biesbroek, “Transformational change: governance interventions for climate change adaptation from a continuous change perspective,” Apr. 03, 2017, *Routledge*. doi: 10.1080/09640568.2016.1168288.

[90] D. Weaver, “Towards Sustainable Mass Tourism: Paradigm Shift or Paradigm Nudge?,” *Tourism Recreation Research*, vol. 32, no. 3, pp. 65–69, 2007, doi: 10.1080/02508281.2007.11081541.

[91] C. J. A. M. Termeer and A. Dewulf, “A small wins framework to overcome the evaluation paradox of governing wicked problems,” *Policy Soc.*, vol. 38, no. 2, pp. 298–314, Apr. 2019, doi: 10.1080/14494035.2018.1497933.

[92] D. Weaver, B. D. Moyle, C. lee McLennan, and L. Casali, “Taming the wicked problem of climate change with ‘virtuous challenges’: An integrated management heuristic,” *J. Environ. Manage.*, vol. 347, Dec. 2023, doi: 10.1016/j.jenvman.2023.119136.

[93] S. Pfenninger, A. Hawkes, and J. Keirstead, “Energy systems modeling for twenty-first century energy challenges,” 2014, *Elsevier Ltd.* doi: 10.1016/j.rser.2014.02.003.

[94] T. Kühne, “What is a Model?,” 2005. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2005/23

[95] V. Krey, “Global energy-climate scenarios and models: a review,” *Wiley Interdiscip. Rev. Energy Environ.*, vol. 3, no. 4, pp. 363–383, 2014.

[96] N. Nakićenović, “Special report on emissions scenarios,” 2000.

[97] M. Beck and T. Krueger, “The epistemic, ethical, and political dimensions of uncertainty in integrated assessment modeling,” Sep. 01, 2016, *Wiley-Blackwell*. doi: 10.1002/wcc.415.

[98] C. W. Churchman, R. L. Ackoff, and E. L. Arnoff, “Introduction to operations research.,” 1957.

[99] P. Wack, “Scenarios: shooting the rapids,” *Harv. Bus. Rev.*, vol. 63, no. 6, pp. 139–150, 1985.

[100] R. D. Shell, “Scenarios: an explorer’s guide,” 2003, *Shell London*.

[101] C. Hepburn, B. O’Callaghan, N. Stern, J. Stiglitz, and D. Zenghelis, “Will COVID-19 fiscal recovery packages accelerate or retard progress on climate change?,” *Oxf. Rev. Econ. Policy*, vol. 36, no. Supplement_1, pp. S359–S381, 2020.

[102] C. S. Holling, “Resilience and stability of ecological systems,” 1973, *International Institute for Applied Systems Analysis Laxenburg*.

[103] B. Walker, C. S. Holling, S. R. Carpenter, and A. Kinzig, “Resilience, adaptability and transformability in social–ecological systems,” *Ecology and society*, vol. 9, no. 2, 2004.

[104] N. M. J. P. Van Beeck, “Classification of Energy Models,” 1999. [Online]. Available: https://research.tilburguniversity.edu/en/publications/6f2cbb5e-2d53-4be6-

[105] P. Mirowski, *Science-mart: privatizing American science*. Harvard University Press, 2011.

[106] R. A. Pielke Jr, *The honest broker: Making sense of science in policy and politics*. Cambridge University Press, 2007.

[107] A. Stirling, “Keep it complex,” *Nature*, vol. 468, no. 7327, pp. 1029–1031, 2010.

[108] B. K. Sovacool, M. Burke, L. Baker, C. K. Kotikalapudi, and H. Wlokas, “New frontiers and conceptual frameworks for energy justice,” *Energy Policy*, vol. 105, pp. 677–691, 2017, doi: 10.1016/j.enpol.2017.03.005.

[109] R. J. Heffron, “Achieving a Just Transition to a Low-Carbon Economy,” 2018.

[110] R. J. Heffron, *The challenge for energy justice*. Springer, 2021.

[111] Silvina. Pugliese *et al.*, *New Zealand Energy Scenarios TIMES-NZ 2.0 : a guide to understanding the TIMES-NZ 2.0 model*. Energy Efficiency and Conservation Authority, 2021.

[112] A. Fattahi, J. Sijm, and A. Faaij, “A systemic approach to analyze integrated energy system modeling tools: A review of national models,” Nov. 01, 2020, *Elsevier Ltd*. doi: 10.1016/j.rser.2020.110195.

[113] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, “On the dangers of stochastic parrots: Can language models be too big?,” in *FAccT 2021 - Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, Inc, Mar. 2021, pp. 610–623. doi: 10.1145/3442188.3445922.

[114] K. Crawford, *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press, 2021.

[115] H. Stachowiak, “Allgemeine modelltheorie,” *(No Title)*, 1973.

[116] M. Morrison and M. S. Morgan, *Models as mediating instruments*. 1999.

[117] J. M. Epstein, “Why Model?,” *Journal of Artificial Societies and Social Simulation*, vol. 11, no. 4, p. 12, 2008.

[118] N. Chomsky, “Language and freedom,” *Resonance*, vol. 4, no. 3, pp. 86–104, 1999.

[119] E. M. Bender and A. Koller, “Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data,” 2020. [Online]. Available: https://www.nytimes.com/2018/11/18/technology/artific

[120] F. Bock and S. Pfenninger-Lee, “Rarely pure and never simple: Exploring perceptions of truth and objectivity in energy modelling and scenarios,” *Energy Res. Soc. Sci.*, vol. 127, no. February, p. 104229, 2025, doi: 10.1016/j.erss.2025.104229.

[121] World Energy Council, “EVOLVING WITH RESILIENCE AND JUSTICE (World_Energy_Trilemma_2024_Full_Report),” 2024.

[122] N. Spittler, G. Gladkykh, A. Diemer, and B. Davidsdottir, “Understanding the current energy paradigm and energy system models for more sustainable energy system development,” Apr. 25, 2019, *MDPI AG*. doi: 10.3390/en12081584.

[123] K. Dyl, A. Kahan, J. Diefenderfer, N. Vincent, B. Cultice, and A. Heisey, “A New Generation of Energy-Economy Modeling at the U.S. Energy Information Administration,” no. October, pp. 1–20, 2024.

[124] E. Winsberg, *Science in the age of computer simulation*. University of Chicago Press, 2019.

[125] F. H. Knight, *Risk, uncertainty and profit*. Рипол Классик, 1965.

[126] B. W. Head, “Wicked Problems in Public Policy Understanding and Responding to Complex Challenges,” 2022.

[127] Plato, *Republic*. 2004.

[128] R. Loulou, U. Remme, A. Kanudia, A. Lehtila, and G. Goldstein, “Documentation for the times model part ii,” *Energy Technology Systems Analysis Programme*, 2005.

[129] J. Barnes, “Posterior analytics,” 1994.

[130] C. G. Heaps, “Long-range Energy Alternatives Planning (LEAP) system.[Software version: 2018.1. 38] Stockholm Environment Institute,” *Somerville, MA, USA*, 2016.

[131] A. A. Long and D. N. Sedley, *The Hellenistic philosophers: Volume 2, Greek and Latin texts with notes and bibliography*. Cambridge University Press, 1987.

[132] P. P. Craig, A. Gadgil, and J. G. Koomey, “WHAT CAN HISTORY TEACH US? A Retrospective,” 2002.

[133] N. Metropolis and S. Ulam, “The monte carlo method,” *J. Am. Stat. Assoc.*, vol. 44, no. 247, pp. 335–341, 1949.

[134] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on stochastic programming: modeling and theory*. SIAM, 2021.

[135] R. N. Langlois and M. M. Cosgel, “FRANK KNIGHT ON RISK, UNCERTAINTY, AND THE FIRM: A NEW INTERPRETATION,” *Econ. Inq.*, vol. 31, no. 3, pp. 456–465, 1993, doi: 10.1111/j.1465-7295.1993.tb01305.x.

[136] M. G. Morgan, M. Henrion, and M. Small, *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge university press, 1990.

[137] W. E. Walker, M. Haasnoot, and J. H. Kwakkel, “Adapt or perish: A review of planning approaches for adaptation under deep uncertainty,” *Sustainability*, vol. 5, no. 3, pp. 955–979, 2013.

[138] N. Spittler, G. Gladkykh, A. Diemer, and B. Davidsdottir, “Understanding the current energy paradigm and energy system models for more sustainable energy system development,” *Energies (Basel).*, vol. 12, no. 8, p. 1584, 2019.

[139] L. G. Fishbone and H. Abilock, “Markal, a linear-programming model for energy systems analysis: Technical description of the bnl version,” *Int. J. Energy Res.*, vol. 5, no. 4, pp. 353–375, 1981.

[140] M. Grubb, J. Edmonds, P. Ten Brink, and M. Morrison, “The costs of limiting fossil-fuel CO2 emissions: a survey and analysis,” *Annual Review of Energy and the environment*, vol. 18, no. 1, pp. 397–478, 1993.

[141] R. Descartes, “Discourse on method and other related writings,” 1999, *New York: Penguin Books (Original work published in 1637)*.

[142] F. Bacon, “1620 [2000] The New Organon,” 2000, *Cambridge, UK, Cambridge, University Press, CXXVI, trad. Lorquet*.

[143] D. Hume, “An enquiry concerning human understanding,” in *Seven masterpieces of philosophy*, Routledge, 2016, pp. 183–276.

[144] T. Bayes, “An essay towards solving a problem in the doctrine of chances,” *Biometrika*, vol. 45, no. 3–4, pp. 296–315, 1958.

[145] M. De Laplace, *A philosophical essay on probabilities*. Courier Corporation, 1995.

[146] A. N. Kolmogorov, *Foundations of the theory of probability: Second English Edition*. Courier Dover Publications, 2018.

[147] J. Von Neumann and O. Morgenstern, “Theory of games and economic behavior: 60th anniversary commemorative edition,” in *Theory of games and economic behavior*, Princeton university press, 2007.

[148] L. J. Savage, *The foundations of statistics*. Courier Corporation, 1972.

[149] P. A. Jensen and J. F. Bard, *Operations research models and methods*. John Wiley & Sons, 2002.

[150] J. Forrester, “W.(1961). Industrial Dynamics,” *Waltham MA, Pegasus Communications*, 1961.

[151] D. H. Meadows, D. L. Meadows, J. Randers, and W. W. Behrens, “‘Perspectives, Problems, and Models’: from The Limits to Growth (1972),” in *The Sustainable Urban Development Reader*, Routledge, 2022, pp. 34–39.

[152] K. Popper, *The logic of scientific discovery*. Routledge, 2005.

[153] T. S. Kuhn, *The structure of scientific revolutions*, vol. 962. University of Chicago press Chicago, 1997.

[154] L. A. Wolsey and G. L. Nemhauser, *Integer and combinatorial optimization*. John Wiley & Sons, 1999.

[155] G. B. Dantzig, “Linear programming and extensions,” 2016.

[156] S. Messner and M. Strubegger, “User’s Guide for MESSAGE III,” 1995.

[157] International Energy Agency (IEA-ETSAP), “TIMES: The Integrated MARKAL-EFOM System Model.” Accessed: Jul. 09, 2024. [Online]. Available: https://iea-etsap.org/index.php/etsap-tools/model-generators/times

[158] T. Brown, J. Hörsch, and D. Schlachtberger, “PyPSA: Python for Power System Analysis,” *J. Open Res. Softw.*, vol. 6, no. 1, p. 4, Jan. 2018, doi: 10.5334/jors.188.

[159] D. H. Meadows, D. L. Meadows, J. Randers, and W. W. Behrens III, “The limits to growth-club of rome,” 1972.

[160] J. M. Reilly, *Global energy: assessing the future*. Oxford University Press, 1985.

[161] A. Fuller, Z. Fan, C. Day, and C. Barlow, “Digital twin: enabling technologies, challenges and open research,” *IEEE access*, vol. 8, pp. 108952–108971, 2020.

[162] J. R. Birge and F. Louveaux, *Introduction to stochastic programming*. Springer, 1997.

[163] H. A. Simon, *Models of man: social and rational; mathematical essays on rational human behavior in society setting*. New York: Wiley, 1957.

[164] K. Deb, “Nonlinear goal programming using multi-objective genetic algorithms,” *Journal of the Operational Research Society*, vol. 52, no. 3, pp. 291–302, 2001.

[165] E.-G. Talbi, *Metaheuristics: from design to implementation*. John Wiley & Sons, 2009.

[166] S. Hajkowicz and A. Higgins, “A comparison of multiple criteria analysis techniques for water resource management,” *Eur. J. Oper. Res.*, vol. 184, no. 1, pp. 255–265, Jan. 2008, doi: 10.1016/j.ejor.2006.10.045.

[167] M. Panteli and P. Mancarella, “The grid: Stronger, bigger, smarter?: Presenting a conceptual framework of power system resilience,” *IEEE Power and Energy Magazine*, vol. 13, no. 3, pp. 58–66, 2015.

[168] S. Hosseini, K. Barker, and J. E. Ramirez-Marquez, “A review of definitions and measures of system resilience,” *Reliab. Eng. Syst. Saf.*, vol. 145, pp. 47–61, 2016.

[169] S. D. Pohekar and M. Ramachandran, “Application of multi-criteria decision making to sustainable energy planning—A review,” *Renewable and sustainable energy reviews*, vol. 8, no. 4, pp. 365–381, 2004.

[170] S. Yates and R. Fellow, “Descartes and Methodological Doubt: Was the Cogito Necessary? version (2.0),” 2002.

[171] D. Hume, “1748. An Enquiry Concerning Human Understanding, ed. Selby-Bigge,” 1893, *Oxford University Press*.

[172] J. Dewey, *Logic-The theory of inquiry*. Read Books Ltd, 2018.

[173] S. 0 Funtowicz and J. R. Ravetz, “SCIENCE FOR THE POST-NORMAL AGE,” 1993.

[174] S. Bankes, “Exploratory Modeling for Policy Analysis,” *Oper. Res.*, vol. 41, no. 3, pp. 435–449, Jun. 1993, doi: 10.1287/opre.41.3.435.

[175] H. M. Regan *et al.*, “Robust decision-making under severe uncertainty for conservation management,” *Ecological Applications*, vol. 15, no. 4, pp. 1471–1477, 2005, doi: 10.1890/03-5419.

[176] J. M. Mulvey, R. J. Vanderbei, and S. A. Zenios, “Robust optimization of large-scale systems,” *Oper. Res.*, vol. 43, no. 2, pp. 264–281, 1995.

[177] M. Workman, G. Darch, K. Dooley, G. Lomax, J. Maltby, and H. Pollitt, “Climate policy decision making in contexts of deep uncertainty-from optimisation to robustness,” *Environ. Sci. Policy*, vol. 120, pp. 127–137, 2021.

[178] EECA, “TIMES-NZ energy model and scenarios,” 2022.

[179] W. E. Walker *et al.*, “Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support,” *Integrated assessment*, vol. 4, no. 1, pp. 5–17, 2003.

[180] Openmod, “Open Energy Modelling Initiative,” 2024. Accessed: Mar. 09, 2024. [Online]. Available: https://wiki.openmod-initiative.org/wiki/Main_Page

[181] M. Winskel, “Beyond the disruption narrative: Varieties and ambiguities of energy system change,” *Energy Res. Soc. Sci.*, vol. 37, pp. 232–237, 2018.

[182] C. Guivarch, R. Lempert, and E. Trutnevyte, “Scenario techniques for energy and environmental research: An overview of recent developments to broaden the capacity to deal with complexity and uncertainty,” *Environmental Modelling and Software*, vol. 97, pp. 201–210, 2017, doi: 10.1016/j.envsoft.2017.07.017.

[183] A. Stirling, “‘Opening up’ and ‘closing down’ power, participation, and pluralism in the social appraisal of technology,” *Sci. Technol. Human Values*, vol. 33, no. 2, pp. 262–294, 2008.

[184] J. D. Sterman, “Business Dynamics: Systems thinking and modeling for a complex world,” *MacGraw-Hill Company*, 2000.

[185] J. Forrester, “Origin of system dynamics,” *System Dynamics Society*, 1961.

[186] J. Ladyman, J. Lambert, and K. Wiesner, “What is a complex system?,” 2013, *Kluwer Academic Publishers*. doi: 10.1007/s13194-012-0056-8.

[187] J. D. Sterman, “SYSTEM DYNAMICS: SYSTEMS THINKING AND MODELING FOR A COMPLEX WORLD.”

[188] R. P. Hämäläinen and T. J. Lahtinen, “Path dependence in Operational Research-How the modeling process can influence the results,” *Operations Research Perspectives*, vol. 3, pp. 14–20, 2016, doi: 10.1016/j.orp.2016.03.001.

[189] D. H. Meadows, “Thinking in systems: A primer,” *Sustainability Institute*, 2008.

[190] A. Martínez-Reyes, J. Lieu, G. de Vries, and T. Hoppe, “Introducing a typology of energy regions: A systematic literature review,” Jan. 01, 2025, *Elsevier Ltd*. doi: 10.1016/j.rser.2024.114961.

[191] D. Li, P. Xu, J. Gu, and Y. Zhu, “A Review of Reliability Research in Regional Integrated Energy System: Indicator, Modeling, and Assessment Methods,” Nov. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/buildings14113428.

[192] J. M. Weinand, F. Scheller, and R. McKenna, “Reviewing energy system modelling of decentralized energy autonomy,” Jul. 15, 2020, *Elsevier Ltd*. doi: 10.1016/j.energy.2020.117817.

[193] S. C. Bhattacharyya and G. R. Timilsina, “A review of energy system models,” Nov. 23, 2010. doi: 10.1108/17506221011092742.

[194] D. Connolly, H. Lund, B. V. Mathiesen, and M. Leahy, “A review of computer tools for analysing the integration of renewable energy into various energy systems,” 2010, *Elsevier Ltd*. doi: 10.1016/j.apenergy.2009.09.026.

[195] M. Ventosa, Á. Baíllo, A. Ramos, and M. Rivier, “Electricity market modeling trends,” *Energy Policy*, vol. 33, no. 7, pp. 897–913, May 2005, doi: 10.1016/j.enpol.2003.10.013.

[196] J. H. Kwakkel, H. S. I. Vreugdenhil, and J. H. Slinger, “Paradigms in sustainability research,” in *ERSCP-EMSU 2010 Conference*, ERSCP, 2010, pp. 1–15.

[197] J. Decarolis *et al.*, “Formalizing best practice for energy system optimization modelling,” 2017.

[198] L. M. H. Hall and A. R. Buckley, “A review of energy systems models in the UK: Prevalent usage and categorisation,” *Appl. Energy*, vol. 169, pp. 607–628, 2016.

[199] S. Pfenninger, “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,” *Appl. Energy*, vol. 197, pp. 1–13, 2017, doi: 10.1016/j.apenergy.2017.03.051.

[200] R. Morrison, “Energy system modeling: Public transparency, scientific reproducibility, and open development,” *Energy Strategy Reviews*, vol. 20, pp. 49–63, 2018.

[201] V. Atias, “Opportunities and Challenges of Using Artificial Intelligence in Energy Communities,” in *International Conference Automatics and Informatics, ICAI 2023 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 508–513. doi: 10.1109/ICAI58806.2023.10339026.

[202] Y. Lin, J. Tang, J. Guo, S. Wu, and Z. Li, “Advancing AI-Enabled Techniques in Energy System Modeling: A Review of Data-Driven, Mechanism-Driven, and Hybrid Modeling Approaches,” *Energies (Basel).*, vol. 18, no. 4, Feb. 2025, doi: 10.3390/en18040845.

[203] J. Beckman, T. Hertel, and W. Tyner, “Validating energy-oriented CGE models,” *Energy Econ.*, vol. 33, no. 5, pp. 799–806, Sep. 2011, doi: 10.1016/j.eneco.2011.01.005.

[204] K. Matsumoto and S. Fujimori, “CGE models in energy economics_25_09_04_13_55_13,” 2019.

[205] K. L. Szeto, “A dynamic computable general equilibrium (CGE) model of the New Zealand economy,” 2002. [Online]. Available: www.treasury.govt.nz

[206] N. Winchester and D. White, “Economics Working Paper Series The Climate PoLicy ANalysis (C-PLAN) Model, Version 1.0,” 2021.

[207] Ganesh. Nana, Kel. Sanderson, and Rob. Hodgson, *Economic impacts of immigration : scenarios using a computable general equilibrium model of the New Zealand economy*. Dept. of Labour, 2009.

[208] N. Robson, “A Multi-Regional Computable General Equilibrium Model for New Zealand,” 2009.

[209] S. Paltsev, “Moving from static to dynamic general equilibrium economic models,” *Notes for a beginner in MPSGE Technical Note (4)*, 2004.

[210] E. Van der Werf, “Production functions for climate policy modeling: An empirical analysis,” *Energy Econ.*, vol. 30, no. 6, pp. 2964–2979, 2008.

[211] C. Böhringer and T. F. Rutherford, “Combining bottom-up and top-down,” *Energy Econ.*, vol. 30, no. 2, pp. 574–596, 2008.

[212] S. S. Scrieciu, “The inherent dangers of using computable general equilibrium models as a single integrated modelling framework for sustainability impact assessment. A critical note on Böhringer and Löschel (2006),” *Ecological Economics*, vol. 60, no. 4, pp. 678–684, 2007.

[213] J. Sathaye and A. H. Sanstad, “Bottom-up energy modeling,” 2004.

[214] International Energy Agency (IEA-ETSAP), “MARKAL: Market Allocation Model,” 2024, Accessed: Jul. 09, 2024. [Online]. Available: https://iea-etsap.org/index.php/etsap-tools/model-generators/markal

[215] R. Loulou and M. Labriet, “ETSAP-TIAM: The TIMES integrated assessment model Part I: Model structure,” *Computational Management Science*, vol. 5, no. 1–2, pp. 7–40, Feb. 2008, doi: 10.1007/s10287-007-0046-z.

[216] S. Pfenninger and B. Pickering, “Calliope: a multi-scale energy systems modelling framework,” *J. Open Source Softw.*, vol. 3, no. 29, p. 825, Sep. 2018, doi: 10.21105/JOSS.00825.

[217] P. A. Østergaard, H. Lund, J. Z. Thellufsen, P. Sorknæs, and B. V. Mathiesen, “Review and validation of EnergyPLAN,” Oct. 01, 2022, *Elsevier Ltd*. doi: 10.1016/j.rser.2022.112724.

[218] S. Lo Piano *et al.*, “Unpacking uncertainty in the modelling process for energy policy making,” 2021.

[219] D. Huppmann *et al.*, “The MESSAGEix Integrated Assessment Model and the ix modeling platform (ixmp): An open framework for integrated and cross-cutting analysis of energy, climate, the environment, and sustainable development,” *Environmental Modelling and Software*, vol. 112, pp. 143–156, Feb. 2019, doi: 10.1016/j.envsoft.2018.11.012.

[220] G. Luderer *et al.*, “Description of the REMIND model (Version 1.6),” 2015.

[221] P. Fortes, S. Simões, J. Seixas, D. Van Regemorter, and F. Ferreira, “Top-down and bottom-up modelling to support low-carbon scenarios: climate policy implications,” *Climate policy*, vol. 13, no. 3, pp. 285–304, 2013.

[222] K. S. Andersen, “Integrated Energy and Macroeconomic Modelling,” 2019.

[223] D. White *et al.*, “Energy-and multi-sector modelling of climate change mitigation in New Zealand: current practice and future needs,” 2018.

[224] P. A. Østergaard, “Reviewing optimisation criteria for energy systems analyses of renewable energy integration,” *Energy*, vol. 34, no. 9, pp. 1236–1245, 2009.

[225] T. Kannengießer *et al.*, “Reducing computational load for mixed integer linear programming: an example for a district and an Island energy system,” *Energies (Basel).*, vol. 12, no. 14, p. 2825, 2019.

[226] C. Sagastizábal, “Divide to conquer: Decomposition methods for energy optimization,” in *Mathematical Programming*, Aug. 2012, pp. 187–222. doi: 10.1007/s10107-012-0570-7.

[227] L. Tesfatsion and K. L. Judd, *Handbook of computational economics: agent-based computational economics*, vol. 2. Elsevier, 2006.

[228] H. K. Ringkjøb, P. M. Haugan, and I. M. Solbrekke, “A review of modelling tools for energy and electricity systems with large shares of variable renewables,” Nov. 01, 2018, *Elsevier Ltd*. doi: 10.1016/j.rser.2018.08.002.

[229] D. Goldsman and P. Goldsman, “Discrete-event simulation,” in *Modeling and Simulation in the Systems Engineering Life Cycle: Core Concepts and Accompanying Lectures*, Springer, 2015, pp. 103–109.

[230] A. Sweetser, “A comparison of system dynamics (SD) and discrete event simulation (DES),” in *17th International Conference of the System Dynamics Society*, 1999, p. 23.

[231] C. R. Paulista, T. A. Peixoto, and J. J. de Assis Rangel, “Modeling and discrete event simulation in industrial systems considering consumption and electrical energy generation,” *J. Clean. Prod.*, vol. 224, pp. 864–880, Jul. 2019, doi: 10.1016/j.jclepro.2019.03.248.

[232] H. Wang *et al.*, “Multi-scale Simulation of Complex Systems: A Perspective of Integrating Knowledge and Data,” *ACM Comput. Surv.*, vol. 56, no. 12, Oct. 2024, doi: 10.1145/3654662.

[233] K. J. ARRow and G. DEBREU, “ECONO MI ETRICA,” *Econometrica*, vol. 22, no. 3, pp. 265–290, 1954.

[234] J. Kornai, “Economic systems theory and general equilibrium theory,” *Acta Oeconomica*, pp. 297–317, 1971.

[235] F. Petri, *General equilibrium, capital and macroeconomics: A key to recent controversies in equilibrium theory*. Edward Elgar Publishing, 2004.

[236] H. Özel, “The Explanatory Role of the General Equilibrium Theory: An Outline into a Critique of Neoclassical Economics,” *Hacettepe Üniversitesi Iktisadi ve Idari Bilimler Fakültesi Dergisi*, vol. 18, no. 1, pp. 257–285, 2000.

[237] C. Fyfe, P. Taptiklis, D. White, and N. Winchester, “Review of emissions data and modelling systems (Phase 1) Report,” Motu Economic and Public Policy Research, Wellington, NZ, Jul. 2023. doi: 10.29310/WP.2023.06.

[238] S. Kerr and C. Leining, “Uncertainty, Risk and Investment and the NZ ETS,” 2019. [Online]. Available: https://ssrn.com/abstract=3477062

[239] M. Chang *et al.*, “Trends in tools and approaches for modelling the energy transition,” *Appl. Energy*, vol. 290, p. 116731, 2021.

[240] S. Pfenninger *et al.*, “Opening the black box of energy modelling: Strategies and lessons learned,” *Energy Strategy Reviews*, vol. 19, pp. 63–71, Jan. 2018, doi: 10.1016/j.esr.2017.12.002.

[241] L. Börjeson, M. Höjer, K. H. Dreborg, T. Ekvall, and G. Finnveden, “Scenario types and techniques: Towards a user’s guide,” *Futures*, vol. 38, no. 7, pp. 723–739, Sep. 2006, doi: 10.1016/j.futures.2005.12.002.

[242] I. Renewable Energy Agency, W. Bank, I. Energy Agency, W. Health Organization, and U. Nations Statistics Division, “Tracking SDG7: The Energy Progress Report 2025,” 2025. [Online]. Available: https://trackingsdg7.esmap.org/downloads.

[243] J. H. Kwakkel, W. E. Walker, and V. A. W. J. Marchau, “From Predictive Modeling to Exploratory Modeling: How to use Non-Predictive Models for Decisionmaking under Deep Uncertainty,” 2010.

[244] P. W. F. van Notten, J. Rotmans, M. B. A. van Asselt, and D. S. Rothman, “An updated scenario typology,” *Futures*, vol. 35, no. 5, pp. 423–443, 2003, doi: 10.1016/S0016-3287(02)00090-3.

[245] UNFCCC, “Report of the Conference of the Parties on its nineteenth session, held in Warsaw from 11 to 23 November 2013,” 2014.

[246] J. F. DeCarolis, K. Hunter, and S. Sreepathi, “The case for repeatable analysis with energy economy optimization models,” *Energy Econ.*, vol. 34, no. 6, pp. 1845–1853, Nov. 2012, doi: 10.1016/j.eneco.2012.07.004.

[247] E. Trutnevyte, “Does cost optimization approximate the real-world energy transition?,” *Energy*, vol. 106, pp. 182–193, Jul. 2016, doi: 10.1016/j.energy.2016.03.038.

[248] R. Amara, “WHAT WE HAVE LEARNED ABOUT FORECASTING AND PLANNING,” 1988.

[249] B. P. Bryant and R. J. Lempert, “Thinking inside the box: A participatory, computer-assisted approach to scenario discovery,” *Technol. Forecast. Soc. Change*, vol. 77, no. 1, pp. 34–49, 2010, doi: 10.1016/j.techfore.2009.08.002.

[250] J. H. Kwakkel and E. Pruyt, “Exploratory Modeling and Analysis, an approach for model-based foresight under deep uncertainty,” *Technol. Forecast. Soc. Change*, vol. 80, no. 3, pp. 419–431, Mar. 2013, doi: 10.1016/j.techfore.2012.10.005.

[251] M. Fodstad *et al.*, “Next frontiers in energy system modelling: A review on challenges and the state of the art,” May 01, 2022, *Elsevier Ltd*. doi: 10.1016/j.rser.2022.112246.

[252] D. P. van Vuuren *et al.*, “A new scenario framework for Climate Change Research: Scenario matrix architecture,” *Clim. Change*, vol. 122, no. 3, pp. 373–386, 2014, doi: 10.1007/s10584-013-0906-1.

[253] J. B. Robinson, “FUTURES UNDER GLASS A recipe for people who hate to predict,” 1990.

[254] J. Quist and P. Vergragt, “Past and future of backcasting: The shift to stakeholder participation and a proposal for a methodological framework,” *Futures*, vol. 38, no. 9, pp. 1027–1045, 2006.

[255] K. H. Dreborg, “Essence of Backcasting,” *Futures*, vol. 28, no. 9, pp. 813–828, 1996.

[256] EUROPEAN COMMISSION, “A Clean Planet for all A European long-term strategic vision for a prosperous, modern, competitive and climate neutral economy,” 2018.

[257] IPCC, “Working Group III contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change Mitigation of Climate Change,” 2022. [Online]. Available: www.ipcc.ch

[258] Climate Change Commission, “Ināia tonu nei: a low emissions future for Aotearoa,” 2022.

[259] S. Hallegatte, A. Shah, R. Lempert, C. Brown, and S. Gill, “Investment Decision Making Under Deep Uncertainty Application to Climate Change,” 2012. [Online]. Available: http://econ.worldbank.org.

[260] C. P. Weaver, R. J. Lempert, C. Brown, J. A. Hall, D. Revell, and D. Sarewitz, “Improving the contribution of climate model information to decision making: The value and demands of robust decision frameworks,” *Wiley Interdiscip. Rev. Clim. Change*, vol. 4, no. 1, pp. 39–60, 2013, doi: 10.1002/wcc.202.

[261] R. J. Lempert, D. G. Groves, S. W. Popper, and S. C. Bankes, “A general, analytic method for generating robust strategies and narrative scenarios,” *Manage. Sci.*, vol. 52, no. 4, pp. 514–528, Apr. 2006, doi: 10.1287/mnsc.1050.0472.

[262] H. R. Maier, J. H. A. Guillaume, H. van Delden, G. A. Riddell, M. Haasnoot, and J. H. Kwakkel, “An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together?,” *Environmental Modelling and Software*, vol. 81, pp. 154–164, Jul. 2016, doi: 10.1016/j.envsoft.2016.03.014.

[263] R. Loulou, A. Lehtilä, A. Kanudia, U. Remme, and G. Goldstein, “Energy Technology Systems Analysis Programme,” 2016. [Online]. Available: https://iea-etsap.org/index.php/documentation

[264] P. H. Li, S. Pye, I. Keppo, M. Jaxa-Rozen, and E. Trutnevyte, “Revealing effective regional decarbonisation measures to limit global temperature increase in uncertain transition scenarios with machine learning techniques,” *Clim. Change*, vol. 176, no. 7, pp. 1–23, 2023, doi: 10.1007/s10584-023-03529-w.

[265] S. Collins, J. P. Deane, and B. Ó Gallachóir, “Adding value to EU energy policy analysis using a multi-model approach with an EU-28 electricity dispatch model,” *Energy*, vol. 130, pp. 433–447, 2017, doi: 10.1016/j.energy.2017.05.010.

[266] A. Weidlich and D. Veit, “A critical survey of agent-based wholesale electricity market models,” *Energy Econ.*, vol. 30, no. 4, pp. 1728–1759, Jul. 2008, doi: 10.1016/j.eneco.2008.01.003.

[267] D. W. Bunn and F. S. Oliveira, “Agent-Based Simulation-An Application to the New Electricity Trading Arrangements of England and Wales,” 2001.

[268] Ștefan Ionescu, C. Delcea, N. Chiriță, and I. Nica, “Exploring the Use of Artificial Intelligence in Agent-Based Modeling Applications: A Bibliometric Study,” *Algorithms*, vol. 17, no. 1, Jan. 2024, doi: 10.3390/a17010021.

[269] P. Hall, “Residual biomass fuel projections for New Zealand-Indicative availability by region and source,” 2017. [Online]. Available: www.scionresearch.com

[270] M. Howells *et al.*, “OSeMOSYS: The Open Source Energy Modeling System. An introduction to its ethos, structure and development.,” *Energy Policy*, vol. 39, no. 10, pp. 5850–5870, Oct. 2011, doi: 10.1016/j.enpol.2011.06.033.

[271] J. Wang, P. Pinson, S. Chatzivasileiadis, M. Panteli, G. Strbac, and V. Terzija, “On Machine Learning-Based Techniques for Future Sustainable and Resilient Energy Systems,” *IEEE Trans. Sustain. Energy*, vol. 14, no. 2, pp. 1230–1243, Apr. 2023, doi: 10.1109/TSTE.2022.3194728.

[272] P. Tomar and V. Grover, “Transforming the Energy Sector: Addressing Key Challenges through Generative AI, Digital Twins, AI, Data Science and Analysis,” *EAI Endorsed Transactions on Energy Web*, vol. 10, pp. 1–5, 2023, doi: 10.4108/EW.4825.

[273] Erdiwansyah, R. Mamat, Syafrizal, M. F. Ghazali, F. Basrawi, and S. M. Rosdi, “Emerging role of generative AI in renewable energy forecasting and system optimization,” Dec. 01, 2025, *Elsevier B.V.* doi: 10.1016/j.scca.2025.100099.

[274] M. M. Forootan, I. Larki, R. Zahedi, and A. Ahmadi, “Machine Learning and Deep Learning in Energy Systems: A Review,” Apr. 01, 2022, *MDPI*. doi: 10.3390/su14084832.

[275] H. Zahedi, *Digital Twins for Green Energy Transition: A Professional Guide to Sustainable Innovation*. 2025. doi: 10.13140/RG.2.2.11087.80803.

[276] M. M. Thwe, A. Stefanov, V. S. Rajkumar, and P. Palensky, “Digital Twins for Power Systems: Review of Current Practices, Requirements, Enabling Technologies, Data Federation, and Challenges,” 2025, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2025.3580055.

[277] A. Rasheed, O. San, and T. Kvamsdal, “Digital twin: Values, challenges and enablers from a modeling perspective,” *IEEE access*, vol. 8, pp. 21980–22012, 2020.

[278] A. Glaws *et al.*, “Designing Future Energy Systems with Generative AI,” 2023, doi: 10.1109/XXX.0000.0000000.

[279] G. Antonesi, T. Cioara, I. Anghel, V. Michalakopoulos, E. Sarmas, and L. Toderean, “From Transformers to Large Language Models: A systema□c review of AI applica□ons in the energy sector towards Agen□c Digital Twins.”

[280] W. B. Powell and B. -Delft, “Foundations and Trends ® in Technology, Information and Operations Management Sequential Decision Analytics and Modeling: Modeling with Python,” 2022.

[281] D. Kahneman and A. Tversky, “Prospect theory: An analysis of decision under risk,” in *Handbook of the fundamentals of financial decision making: Part I*, World Scientific, 2013, pp. 99–127.

[282] R. A. Howard, “Decision analysis: Applied decision theory,” 1966.

[283] W. Edwards, “The theory of decision making.,” *Psychol. Bull.*, vol. 51, no. 4, p. 380, 1954.

[284] D. H. Jonassen, “Designing for decision making,” *Educational technology research and development*, vol. 60, no. 2, pp. 341–359, 2012.

[285] A. Church, “ET Bell, Men of Mathematics,” *J. Symbolic Logic*, vol. 2, no. 2, 1937.

[286] I. Hacking, *An introduction to probability and inductive logic*. Cambridge university press, 2001.

[287] P. F. Pelz, P. Groche, M. E. Pfetsch, and M. Schaeffner, *Mastering Uncertainty in Mechanical Engineering*. 2021. [Online]. Available: http://www.springer.

[288] A. Der Kiureghian and O. Ditlevsen, “Aleatory or epistemic? Does it matter?,” *Structural Safety*, vol. 31, no. 2, pp. 105–112, Mar. 2009, doi: 10.1016/j.strusafe.2008.06.020.

[289] S. Buğra Selçuklu, D. W. Coit, and F. A. Felder, “A Classification of Aleatory and Epistemic Uncertainties in Generation Expansion Planning,” 2020.

[290] J. C. HELTON and D. E. BURMASTER, “Treatment of aleatory and epistemic uncertainty,” *Reliability engineering & systems safety*, vol. 54, no. 2–3, 1996.

[291] J. C. Refsgaard, J. P. van der Sluijs, A. L. Højberg, and P. A. Vanrolleghem, “Uncertainty in the environmental modelling process–a framework and guidance,” *Environmental modelling & software*, vol. 22, no. 11, pp. 1543–1556, 2007.

[292] L. Uusitalo, A. Lehikoinen, I. Helle, and K. Myrberg, “An overview of methods to evaluate uncertainty of deterministic models in decision support,” Jan. 01, 2015, *Elsevier Ltd*. doi: 10.1016/j.envsoft.2014.09.017.

[293] A. Krook-Riekkola, C. Berg, E. O. Ahlgren, and P. Söderholm, “Challenges in top-down and bottom-up soft-linking: Lessons from linking a Swedish energy system model with a CGE model,” *Energy*, vol. 141, pp. 803–817, 2017.

[294] J. B. Samon, J. M. Ndomou, and D. H. Fotsa, “Analysis of coupling methods: A review,” 2022.

[295] Siddiqui and Sauleh A, “Iterative methods for passing duals between modules,” 2024. [Online]. Available: www.eia.gov

[296] F. Ding and T. Chen, “On iterative solutions of general coupled matrix equations,” *SIAM J. Control Optim.*, vol. 44, no. 6, pp. 2269–2284, 2006.

[297] S. Pfenninger *et al.*, “Opening the black box of energy modelling: Strategies and lessons learned,” *Energy Strategy Reviews*, vol. 19, pp. 63–71, 2018.

[298] J. Shortridge, T. Aven, and S. Guikema, “Risk assessment under deep uncertainty: A methodological comparison,” *Reliab. Eng. Syst. Saf.*, vol. 159, pp. 12–23, 2017.

[299] M. G. Morgan, M. Henrion, and M. Small, *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge university press, 1990.

[300] A. O’Hagan *et al.*, “Uncertain judgements: eliciting experts’ probabilities,” 2006.

[301] V. Stojkoski, T. Sandev, L. Basnarkov, L. Kocarev, and R. Metzler, “Generalised geometric Brownian motion: Theory and applications to option pricing,” *Entropy*, vol. 22, no. 12, p. 1432, 2020.

[302] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[303] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts, 2018.

[304] R. B. Nelsen, *An introduction to copulas*. Springer, 2006.

[305] P. Embrechts, A. McNeil, and D. Straumann, “Correlation and dependence in risk management: properties and pitfalls,” *Risk management: value at risk and beyond*, vol. 1, pp. 176–223, 2002.

[306] M. D. McKay, R. J. Beckman, and W. J. Conover, “A comparison of three methods for selecting values of input variables in the analysis of output from a computer code,” *Technometrics*, vol. 42, no. 1, pp. 55–61, 2000.

[307] A. B. Owen, “Scrambling Sobol’and Niederreiter–Xing Points,” *J. Complex.*, vol. 14, no. 4, pp. 466–489, 1998.

[308] M. Pinsky and S. Karlin, *An introduction to stochastic modeling*. Academic press, 2010.

[309] R. T. Rockafellar and S. Uryasev, “Optimization of conditional value-at-risk,” *Journal of risk*, vol. 2, pp. 21–42, 2000.

[310] B. Mandelbrot, “The variation of certain speculative prices,” *Journal of business*, vol. 36, no. 4, p. 394, 1963.

[311] X. Yue, S. Pye, J. DeCarolis, F. G. N. Li, F. Rogan, and B. Ó. Gallachóir, “A review of approaches to uncertainty assessment in energy system optimization models,” *Energy strategy reviews*, vol. 21, pp. 204–217, 2018.

[312] M. Paredes-Vergara, J. L. Cerda-Arias, R. Palma-Behnke, and J. Haas, “A framework for integrating deep uncertainty in power systems planning: an application to the Chilean case,” *Appl. Energy*, vol. 398, p. 126437, 2025.

[313] L. J. Savage, “The theory of statistical decision,” *J. Am. Stat. Assoc.*, vol. 46, no. 253, pp. 55–67, 1951.

[314] J. D. Herman, P. M. Reed, H. B. Zeff, and G. W. Characklis, “How should robustness be defined for water systems planning under change?,” *J. Water Resour. Plan. Manag.*, vol. 141, no. 10, p. 04015012, 2015.

[315] J. H. Kwakkel, M. Haasnoot, and W. E. Walker, “Developing dynamic adaptive policy pathways: a computer-assisted approach for developing adaptive strategies for a deeply uncertain world,” *Clim. Change*, vol. 132, no. 3, pp. 373–386, Oct. 2015, doi: 10.1007/s10584-014-1210-4.

[316] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, “Planning and acting in partially observable stochastic domains,” *Artif. Intell.*, vol. 101, no. 1–2, pp. 99–134, 1998.

[317] I. Halperin, *Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions: by Warren B. Powell (ed.), Wiley (2022). Hardback. ISBN 9781119815051.*, vol. 22, no. 12. Taylor & Francis, 2022.

[318] M. T. J. Spaan, “Partially observable Markov decision processes,” in *Reinforcement learning: State-of-the-art*, Springer, 2012, pp. 387–414.

[319] G. E. Monahan, “State of the art—a survey of partially observable Markov decision processes: theory, models, and algorithms,” *Manage. Sci.*, vol. 28, no. 1, pp. 1–16, 1982.

[320] F. Garcia and E. Rachelson, “Markov decision processes,” *Markov Decision Processes in Artificial Intelligence*, pp. 1–38, 2013.

[321] J. Clifton and E. Laber, “Q-learning: Theory and applications,” *Annu. Rev. Stat. Appl.*, vol. 7, no. 1, pp. 279–301, 2020.

[322] R. S. Sutton, D. McAllester, S. Singh, and Y. Mansour, “Policy gradient methods for reinforcement learning with function approximation,” *Adv. Neural Inf. Process. Syst.*, vol. 12, 1999.

[323] F. V Jensen and T. D. Nielsen, *Bayesian Networks and Decision Graphs: February 8, 2007*. Springer, 2007.

[324] A. Shih, A. Choi, and A. Darwiche, “Compiling Bayesian network classifiers into decision graphs,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, pp. 7966–7974.

[325] M. S. Mahbub, M. Cozzini, P. A. Østergaard, and F. Alberti, “Combining multi-objective evolutionary algorithms and descriptive analytical modelling in energy scenario design,” *Appl. Energy*, vol. 164, pp. 140–151, Feb. 2016, doi: 10.1016/j.apenergy.2015.11.042.

[326] J. Boston and J. Lawrence, “Funding climate change adaptation: The case for a new policy framework,” *Policy Quarterly*, vol. 14, no. 2, 2018.

[327] I. P. Holman, C. Brown, T. R. Carter, P. A. Harrison, and M. Rounsevell, “Improving the representation of adaptation in climate change impact models,” *Reg. Environ. Change*, vol. 19, no. 3, pp. 711–721, Mar. 2019, doi: 10.1007/s10113-018-1328-4.

[328] K. Bosomworth and E. Gaillard, “Engaging with uncertainty and ambiguity through participatory ‘Adaptive Pathways’ approaches: Scoping the literature,” *Environmental Research Letters*, vol. 14, no. 9, p. 093007, 2019.

[329] P. M. Reed, D. Hadka, J. D. Herman, J. R. Kasprzyk, and J. B. Kollat, “Evolutionary multiobjective optimization in water resources: The past, present, and future,” *Adv. Water Resour.*, vol. 51, pp. 438–456, 2013.

[330] Z. Xu, *Uncertain multi-attribute decision making: Methods and applications*. Springer, 2015.

[331] J. C. Chacon-Hurtado and L. Scholten, “Decisi-o-rama: An open-source Python library for multi-attribute value/utility decision analysis,” *Environmental Modelling & Software*, vol. 135, p. 104890, 2021.

[332] M. A. Cuesta, T. Castillo-Calzadilla, and C. E. Borges, “A critical analysis on hybrid renewable energy modeling tools: An emerging opportunity to include social indicators to optimise systems in small communities,” *Renewable and Sustainable Energy Reviews*, vol. 122, p. 109691, 2020.

[333] Y. Ben-Haim, *Info-gap decision theory: decisions under severe uncertainty*. Elsevier, 2006.

[334] Y. Ben-Haim, *Info-gap economics: an operational introduction*. Springer, 2010.

[335] M. Sniedovich, “An important maximin theorem: Comments on ‘Decision Making in Times of Knightian Uncertainty: An Info-Gap Perspective,’” 2015.

[336] L. O. Mearns, C. E. Forest, H. J. Fowler, R. Lempert, and R. L. Wilby Editors, “Uncertainty in Climate Change Research,” 2025.

[337] A. K. Dixit and R. S. Pindyck, *Investment under uncertainty*. Princeton university press, 1994.

[338] R. De Neufville and S. Scholtes, *Flexibility in engineering design*. MIT press, 2011.

[339] J. H. Friedman, N. Fisher, and N. I. Fisher, “Bump hunting in high-dimensional data,” 1999. [Online]. Available: https://www.researchgate.net/publication/283550136

[340] L. Breiman, J. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. Chapman and Hall/CRC, 2017.

[341] S. W. Popper, “Robust decision making and scenario discovery in the absence of formal models,” *Futures and Foresight Science*, vol. 1, no. 3–4, Sep. 2019, doi: 10.1002/ffo2.22.

[342] V. Belton and T. Stewart, *Multiple criteria decision analysis: an integrated approach*. Springer Science & Business Media, 2012.

[343] C. C. Price, J. Zhu, and F. S. Hillier, *Multiple Criteria Decision Analysis*. 2016. [Online]. Available: http://www.springer.com/series/6161

[344] M. Cinelli, S. R. Coles, and K. Kirwan, “Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment,” 2014, *Elsevier*. doi: 10.1016/j.ecolind.2014.06.011.

[345] J. M. Keisler, Z. A. Collier, E. Chu, N. Sinatra, and I. Linkov, “Value of information analysis: the state of application,” *Environ. Syst. Decis.*, vol. 34, no. 1, pp. 3–23, 2014.

[346] V. Balali, B. Zahraie, and A. Roozbahani, “Integration of ELECTRE III and PROMETHEE II decision-making methods with an interval approach: Application in selection of appropriate structural systems,” *Journal of computing in civil engineering*, vol. 28, no. 2, pp. 297–314, 2014.

[347] S. Greco, B. Matarazzo, and R. Slowinski, “20. Dominance-Based Rough Set Approach to Knowledge Discovery (I): General Perspective,” 2004.

[348] J. J. Wang, Y. Y. Jing, C. F. Zhang, and J. H. Zhao, “Review on multi-criteria decision analysis aid in sustainable energy decision-making,” Dec. 2009. doi: 10.1016/j.rser.2009.06.021.

[349] J. R. Figueira, V. Mousseau, and B. Roy, “ELECTRE methods,” in *Multiple criteria decision analysis: State of the art surveys*, Springer, 2005, pp. 155–185.

[350] W. Edwards and F. H. Barron, “SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement,” *Organ. Behav. Hum. Decis. Process.*, vol. 60, no. 3, pp. 306–325, 1994.

[351] C. A. Bana e Costa and J.-C. Vansnick, “General overview of the MACBETH approach,” *Advances in multicriteria analysis*, pp. 93–100, 1995.

[352] T. L. Saaty, “The analytic hierarchy process: A new approach to deal with fuzziness in architecture,” *Archit. Sci. Rev.*, vol. 25, no. 3, pp. 64–69, 1982.

[353] R. Lahdelma and P. Salminen, “SMAA-2: Stochastic multicriteria acceptability analysis for group decision making,” *Oper. Res.*, vol. 49, no. 3, pp. 444–454, 2001.

[354] E. Triantaphyllou, “Multi-criteria decision making methods,” in *Multi-criteria decision making methods: A comparative study*, Springer, 2000, pp. 5–21.

[355] C. Rothery *et al.*, “Value of Information Analytical Methods: Report 2 of the ISPOR Value of Information Analysis Emerging Good Practices Task Force,” *Value in Health*, vol. 23, no. 3, pp. 277–286, Mar. 2020, doi: 10.1016/j.jval.2020.01.004.

[356] H. Tuffaha, “Value of information analysis: Are we there yet?,” *PharmacoEconomics-Open*, vol. 5, no. 2, pp. 139–141, 2021.

[357] A. Scolobig and L. Gallagher, “Understanding, analysing and addressing conflicts in co-production,” in *The palgrave handbook of Co-production of public services and outcomes*, Springer, 2020, pp. 613–636.

[358] E. Loeffler and T. Bovaird, *The Palgrave Handbook of Co-Production of Public Services and Outcomes*. 2021.

[359] B. Davidsdottir *et al.*, “Integrated Energy Systems Modeling with Multi-Criteria Decision Analysis and Stakeholder Engagement for Identifying a Sustainable Energy Transition,” *Energies (Basel).*, vol. 17, no. 17, Sep. 2024, doi: 10.3390/en17174266.

[360] R. F. Naill, “A system dynamics model for national energy policy planning,” *Syst. Dyn. Rev.*, vol. 8, no. 1, pp. 1–19, 1992.

[361] A. Ford, *Modeling the environment*, vol. 488. Island press Washington, DC, 2010.

[362] Y. Barlas, "Formal aspects of model validity and validation in system dynamics," *System Dynamics Review: The Journal of the System Dynamics Society*, vol. 12, no. 3, pp. 183–210, 1996.

[363] S. De Marchi and S. E. Page, "Agent-based models," *Annual Review of Political Science*, vol. 17, pp. 1–20, 2014, doi: 10.1146/annurev-polisci-080812-191558.

[364] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proceedings of the national academy of sciences*, vol. 99, no. suppl_3, pp. 7280–7287, 2002.

[365] K. E. Train, *Discrete choice methods with simulation*. Cambridge university press, 2009.

[366] F. M. Bass, "A new product growth for model consumer durables," *Manage. Sci.*, vol. 15, no. 5, pp. 215–227, 1969.

[367] G. Fagiolo, A. Moneta, and P. Windrum, "A critical guide to empirical validation of agent-based models in economics: Methodologies, procedures, and open problems," *Comput. Econ.*, vol. 30, no. 3, pp. 195–226, 2007.

[368] M. Berg, B. Hartley, and O. Richters, "A stock-flow consistent input-output model with applications to energy price shocks, interest rates, and heat emissions," *New J. Phys.*, vol. 17, Jan. 2015, doi: 10.1088/1367-2630/17/1/015011.

[369] X. Xing, K. Guo, D. Zhang, and Q. Ji, "On the interactive effects of climate policies: Insights from a stock-flow consistent model," *Appl. Energy*, vol. 358, Mar. 2024, doi: 10.1016/j.apenergy.2024.122664.

[370] N. Dunz, A. Naqvi, and I. Monasterolo, "Climate sentiments, transition risk, and financial stability in a stock-flow consistent model," *Journal of Financial Stability*, vol. 54, Jun. 2021, doi: 10.1016/j.jfs.2021.100872.

[371] F. Krausmann, D. Wiedenhofer, and H. Haberl, "Growing stocks of buildings, infrastructures and machinery as key challenge for compliance with climate targets," *Global Environmental Change*, vol. 61, Mar. 2020, doi: 10.1016/j.gloenvcha.2020.102034.

[372] G. Wiest, N. Nolzen, F. Baader, A. Bardow, and S. Moret, "Low-regret Strategies for Energy Systems Planning in a Highly Uncertain Future," May 2025, [Online]. Available: http://arxiv.org/abs/2505.13277

[373] P. Industries, *Information on dairy farming systems*, vol. 6, no. May 2023. 2024.

[374] Fonterra, "Edendale's new electrode boiler delivers 20% emissions reduction for site." Accessed: Nov. 25, 2025. [Online]. Available: https://www.fonterra.com/nz/en/our-stories/articles/edendales-new-electrode-boiler-delivers-20-percent-emissions-reduction-for-site.html?utm_source=chatgpt.com

[375] Kane Morison, "Southland Electrical Network Supplementary Information-Substation Load Characteristics," 2023.

[376] Sudesh Kissun, "Fonterra announces new $150m UHT plant for Edendale site." Accessed: Nov. 26, 2025. [Online]. Available: https://www.ruralnewsgroup.co.nz/dairy-news/dairy-general-news/fonterra-announces-new-150m-uht-plant-for-edendale-site?utm_source=chatgpt.com

[377] Karen Trebilcock, “A milestone at Edendale - Dairy Country.” Accessed: Nov. 26, 2025. [Online]. Available: https://dairycountry.co.nz/a-milestone-at-edendale/

[378] Southland Regional Development Agency, “Southland Regional Development Agency 1 Southland Greenhouse Gas Emissions Inventory for 2022,” 2023.

[379] NIWA, “Obtaining climate data from NIWA | Earth Sciences New Zealand | NIWA.” Accessed: Nov. 27, 2025. [Online]. Available: https://niwa.co.nz/climate-and-weather/obtaining-climate-data-niwa

Appendix

Appendix A: Operational definitions of key terms

This appendix defines key terms used in their operational sense within the thesis and the PoC implementation. The intent is to reduce ambiguity and ensure that terms map cleanly onto implemented artefacts, variables, and validation checks.

Artefact: A versioned, file-based data product with a declared schema and temporal/spatial index that serves as the interface contract between modules (e.g., DemandPack, SignalsPack).

Thin-waist contract: A deliberately minimal, stable set of artefacts and fields that mediate coupling between modules, enabling internal refinement without breaking interoperability.

DemandPack: The site demand artefact that provides time-aligned hourly heat (and where applicable electricity) demand traces consistent with annual totals and declared shaping assumptions.

SignalsPack: The regional-interface artefact providing exogenous hourly signals and discrete option menus required for screening (e.g., headroom, tariff, upgrade options).

Epoch / epoch tag: A discrete time-slice used to represent a decision stage or system state (e.g., 2020 baseline; 2024/2027 committed electrification; 2035 pathways).

Pathway: A discrete site transition decision (or bundle of decisions) that defines the utility portfolio and operating regime to be evaluated for a given epoch.

One-pass coupling: A coupling discipline in which the site module exports a coupling artefact (e.g., ΔE_t) that is screened by the regional module without any feedback loop or second site re-optimisation pass.

Feasibility (site ledger): An observable outcome in which the site dispatch meets demand with the declared portfolio and constraints; any residual shortfall is recorded explicitly as an unserved term rather than being hidden in balancing variables.

Hosting capacity / headroom: The time-varying capacity margin at the GXP/interface available to serve incremental electrified demand, represented as an hourly envelope H_t.

Exceedance: The amount by which incremental electricity demand exceeds available headroom (after any assumed upgrade), used as a stress and feasibility indicator in screening.

Shedding: A proxy representation of unserved electricity at the interface level, used to monetise exceedance under VOLL within the robustness screen.

Robustness (RDM sense): Performance evaluated over a futures ensemble Ω, with robustness measured via satisficing rates and regret rather than reliance on a single predicted future.

Appendix B: DemandPack diagnostics for non-illustrated epochs

Chapter 8 illustrates DemandPack diagnostics for 2020. For completeness, corresponding diagnostics for 2025, 2028, and 2035 are included here, supporting discussion of seasonality, peakiness, and allocation consistency.

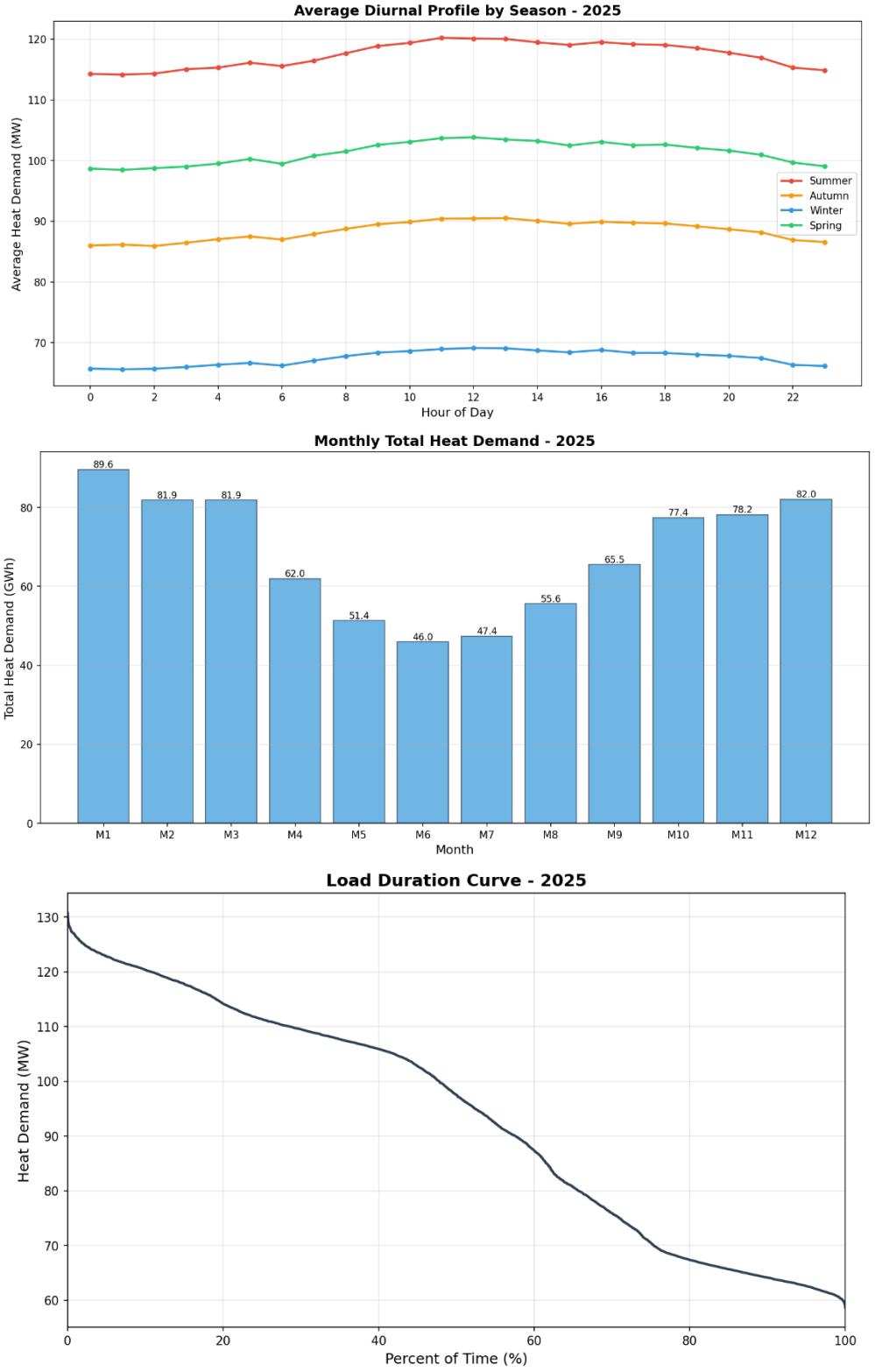


Figure B.1: DemandPack diagnostics for epoch 2025 (hourly profile and duration curve)

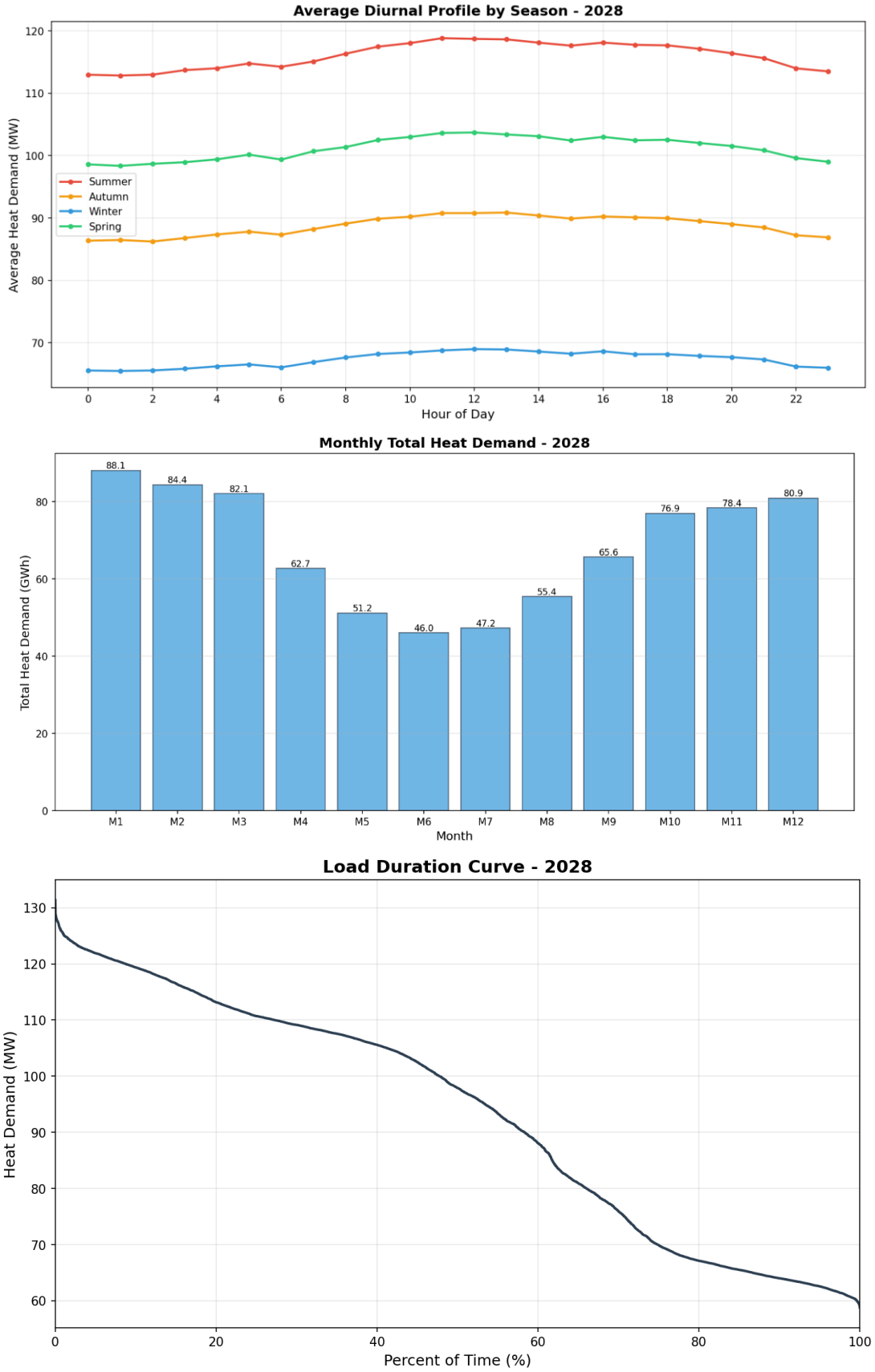


Figure B.2: DemandPack diagnostics for epoch 2028 (hourly profile and duration curve)

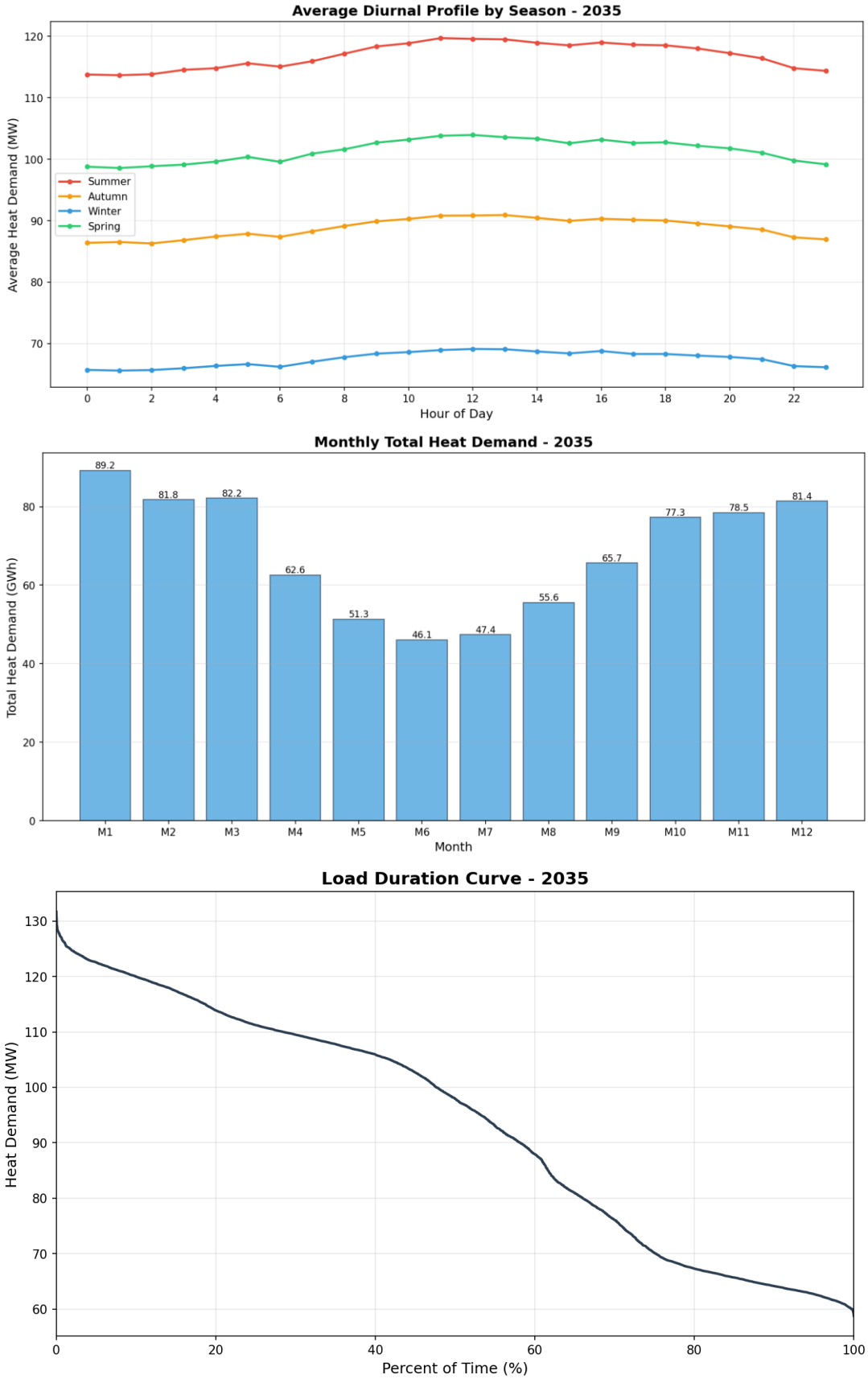


Figure B.3: DemandPack diagnostics for epoch 2035 (hourly profile and duration curve)

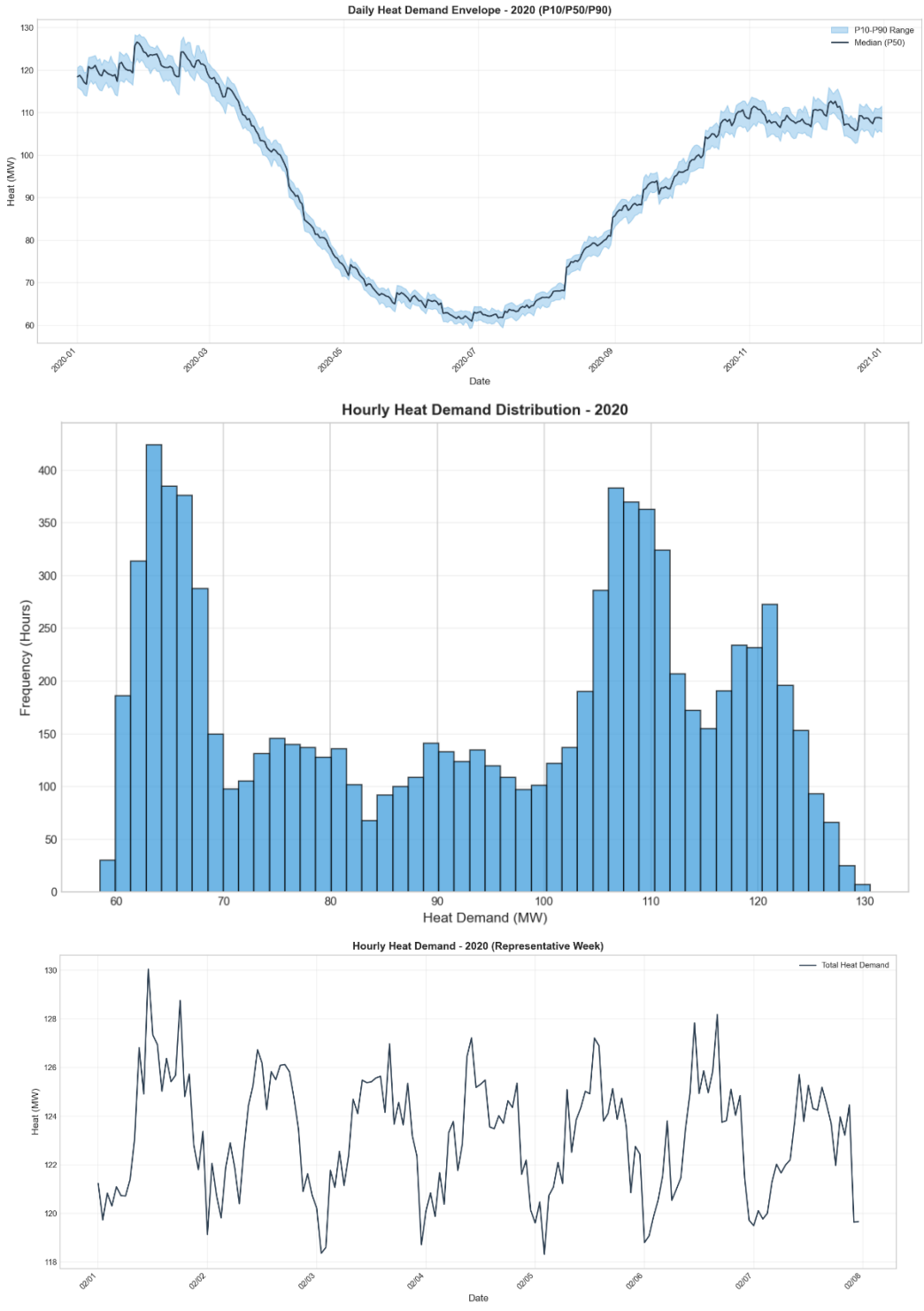


Figure B.4: DemandPack diagnostics for epoch 2020

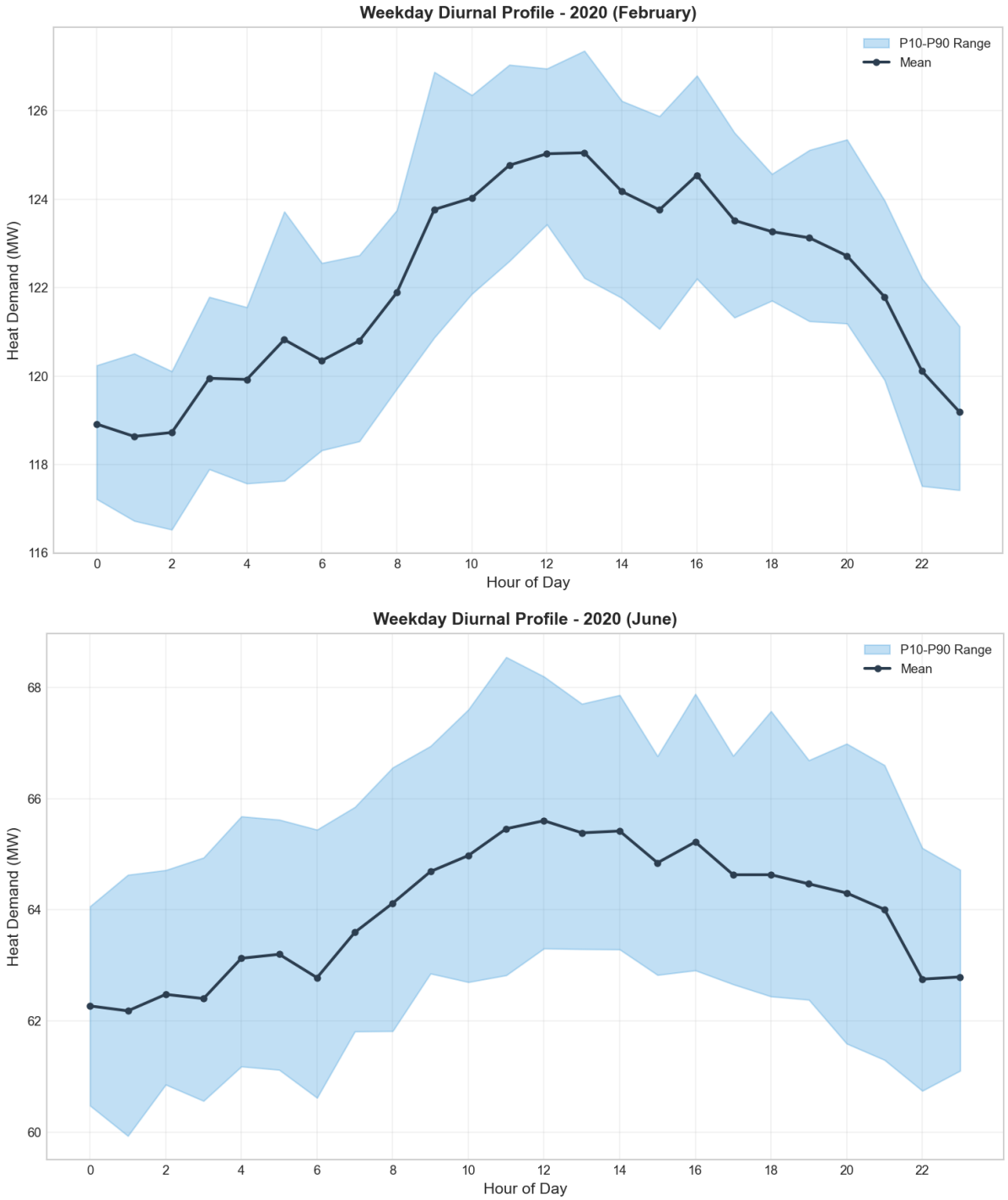


Figure B.5: DemandPack diagnostics for epoch 2020 (weekday diurnal profile for represented months)

Appendix C: Data provenance and source catalogue (PoC)

This appendix records the high-leverage data sources used for the PoC and maps each source onto the artefacts and parameters it informs. The aim is transparency and reproducibility: readers can trace each PoC input back to a documented evidence base, and future work can replace proxies without changing interfaces.

Source	Key quantities used	Role in PoC	Thesis refs
Southland RETA network assessment	Upgrade menu (capacity increments, indicative capex) for Edendale connection options; screening-level basis for $\mathbb{P}$_grid.	SignalsPack: upgrade menu parameters; screening capex.	[40], [369]
Southland RETA biomass assessment	Regional biomass availability (t/TJ) and delivered cost ranges; informs BB scenario environments.	Fuel cost parameters; biomass availability scenarios.	[40]
EECA Regional Heat Demand Database (RHDD)	Fuel shares, regional process-heat demand and emissions context; supports NZ evidence base and parameter plausibility.	Contextual benchmarking; potential priors for annual heat.	[7]
Fonterra / EECA / GIDI public disclosures	Committed electrification investments and indicative abatement; bounds portfolio evolution across epochs.	Site portfolio configuration (2024, 2027 committed additions).	[15], [37], [63]
MBIE energy and climate publications	National policy and sector statistics (process heat demand, fuel mix trends) used for narrative and cross-checking.	Literature/evidence base and contextual parameters.	[6], [67]
Transpower reports on process heat electrification	System and network considerations shaping feasibility and timing of electrification.	Context for regional constraints and lead times.	[61]

Appendix D: Baseline parameters of the proof of concept

This appendix lists the baseline values used to produce the results in Chapter 8, so the reported results can be read and checked without access to the repository. Values are as recorded in the input files of the thesis-era release (Process-Heat-RDM, tag poc_20260105_release02_freeze, commit 368015d). Sources are catalogued in Appendix C. All monetary values are in NZD (2023). Parameters marked as PoC assumptions are screening values chosen to demonstrate the framework; they are not calibrated estimates.

D.1 Site demand and utility portfolio

Annual heat demand is held at 819 GWh in every epoch, anchored to the RHDD record for the Edendale site [7]. Differences between epochs therefore arise only from the declared utility portfolio (Table D.1).

Table D.1: Utility portfolio by epoch

Unit	Technology (carrier)	Capacity (MW_th)	Efficiency η (LHV)	Emission factor (tCO_2/MWh_fuel)	2020	2025	2028	2035_EB	2035_BB
CB1	Coal boiler (lignite)	26	0.78	0.25	✓	–	–	–	–
CB2	Coal boiler (lignite)	25	0.80	0.25	✓	✓	–	–	–
CB3	Coal boiler (lignite)	40	0.86	0.25	✓	✓	–	–	–
CB4	Coal boiler (lignite)	40	0.88	0.25	✓	✓	✓	–	–
EB1	Electrode boiler (electricity)	30	0.99	0	–	✓	✓	✓	✓
EB2	Electrode boiler (electricity)	35	0.99	0	–	–	✓	✓	✓
EB3	Electrode boiler (electricity)	35	0.99	0	–	–	✓	✓	✓
EB4	Electrode boiler (electricity)	40	0.99	0	–	–	–	✓	–
BB1	Biomass boiler (biomass)	40	0.85	0.0059	–	–	–	–	✓
Installed capacity					**131**	**135**	**140**	**140**	**140**

Notes:

- Coal-unit capacities follow the RHDD site record [7], which differs from original nameplate ratings.
- EB1–EB3 represent the committed electrification investments [15], [37], [63]. EB4 and BB1 are the candidate 2035 units that define the two pathways. Their capacities and efficiencies are PoC assumptions.
- Electrode boilers carry no site emissions; electricity-sector emissions lie outside the site boundary (Section 7.4.2).
- The biomass factor is a supply-chain proxy; stack CO_2 from biomass is treated as neutral.
- Unit availability is set by the epoch maintenance windows (Section 7.5.2).

D.2 Signals, emission factors and grid-screening parameters

Table D.2: Baseline signals and screening parameters

Panel A: Site-consumed signals by epoch

Parameter	2020	2025	2028	2035	Unit	Basis
Electricity tariff (flat)	105	145	130	110	NZD/MWh_e	PoC assumption
Coal price	7.0	7.0	7.0	10.0	NZD/MWh_fuel	PoC assumption
Biomass price	35	35	38	42	NZD/MWh_fuel	PoC assumption, informed by RETA biomass assessment [40]
ETS carbon price	30	60	70	90	NZD/tCO_2	PoC assumption
Unserved-heat penalty (site)	20,000	20,000	20,000	20,000	NZD/MWh_th	Feasibility diagnostic

Panel B: Grid-interface screening

Parameter	Value	Basis
GXP headroom H_t	Hourly series, 30–76 MW (median ≈ 52 MW)	PoC-generated envelope (Section 8.4; Table 8.1)
S0 No upgrade	0 MW; NZD 0	–
S1 Upgrade option 1	+21 MW (N); capex NZD 9.6 m; lead time 42 months	RETA Southland network assessment [40], [369]
S2 Upgrade option 2	+32 MW (N–1); capex NZD 34.1 m; lead time 42 months	RETA Southland network assessment [40], [369]
S3 Upgrade option 3	+97 MW (N); capex NZD 55.9 m; lead time 48 months	RETA Southland network assessment [40], [369]
S4 Upgrade option 4	+150 MW (N); capex NZD 75.0 m; lead time 60 months	PoC extension beyond the RETA menu
Real discount rate; asset life	7%; 40 years (CRF = 0.0750)	Screening assumption
Consenting uplift	Reference 1.25; sampled per future (Table D.3)	Screening assumption
Value of lost load (interface)	Sampled per future (Table D.3)	Screening assumption

D.3 Uncertainty sampling bounds

The 100 paired futures are fixed in a futures table shared by both pathways (Section 7.6.3). Table D.3 gives the bounds within which each driver was varied; the values realised across the 100 futures are reported in Table 8.8.

Table D.3: Sampling bounds of the uncertain drivers

Driver	Applies to	Bounds or options	Distribution
Headroom multiplier U_head	Grid screen	0.70–1.00	Fixed futures table
Incremental-load multiplier U_inc	Grid screen	0.85–1.15	Fixed futures table
Upgrade capex multiplier U_capex	Grid screen	0.80–1.50	Fixed futures table
Consenting uplift U_cons	Grid screen	1.00–1.60	Fixed futures table
Value of lost load U_VOLL	Grid screen	{5,000; 10,000; 15,000; 20,000} NZD/MWh	Discrete
Electricity price multiplier P_elec	Overlay	0.65–1.35 (mode 1.0)	Triangular, seed 42
Biomass price multiplier P_bio	Overlay	0.90–2.50 (mode 1.2)	Triangular, seed 42
ETS price multiplier P_ETS	Overlay	0.80–2.00 (mode 1.0)	Triangular, seed 42

D.4 DemandPack construction

The synthetic hourly profile (Section 7.5.1) distributes the annual 819 GWh using the settings in Table D.4. The factor tables themselves are archived with the release.

Table D.4: DemandPack shaping settings (all epochs)

Component	Setting
Seasonal shape	Monthly anchor factors (anchored on day 15), time-interpolated, 21-day smoothing, wrapped across the year
Weekday shape	Day-of-week scaling factors
Within-day shape	Hour-of-day factors
Weekly drift	AR(1), σ = 0.03, ρ = 0.9, bounded 0.9–1.1
Hourly noise	σ = 0.01, seed 42
Chronology	Hourly UTC; 8,784 hours in 2020 (leap year), 8,760 otherwise
Operating basis	24 h/day, 350 days/year

D.5 Reproduction check

Two results in Table 8.2 can be recomputed directly from Tables D.1–D.2.

- **2035_EB electricity cost.** All heat is supplied by electrode boilers: 819 GWh ÷ 0.99 = 827.3 GWh of electricity. At NZD 110/MWh this is NZD 91.0 m, matching Tables 8.2 and 8.7.
- **2020 emissions.** Under proportional dispatch, each coal unit supplies heat in proportion to its capacity (26 : 25 : 40 : 40 of 131 MW). Fuel input is therefore $\Sigma\ (819 \times K_u/131)/\eta_u \approx$ 978.8 GWh. At 0.25 tCO_2/MWh_fuel this gives 244.7 kt, matching Table 8.2. With coal at NZD 7.0/MWh and ETS at NZD 30/t, the annual cost is ≈ NZD 6.9 m + 7.3 m = 14.2 m, also matching Table 8.2.

Appendix E: Code and data availability

The code, input files and configuration used to produce the results reported in Chapters 7 and 8 are openly available in the proof-of-concept repository:

- Repository: Process-Heat-RDM, https://github.com/Ahmad-Mahmoudi-coder/Process-Heat-RDM (MIT licence).
- Thesis state: tag *thesis-poc-canonical* (also tagged *poc_20260105_release02_freeze*), commit 368015d. This is the frozen release that produced result bundle poc_20260105_release02, from which every figure and table in Chapter 8 is drawn, and the source of the baseline values in Appendix D.
- Reproduction layer: tag *thesis-poc-reproduction*, commit 568e6cc, which adds the scripts used to regenerate the Chapter 8 figures; the model inputs are unchanged from the thesis state.
- Archived release: Process-Heat-RDM: Proof of Concept for Decision-Centred Modelling (DCM), v1.0.3, Zenodo, 2026, https://doi.org/10.5281/zenodo.21800568. This is a later packaged release of the same repository; the tags above identify the exact state used in this thesis.

The GXP headroom series used in the grid screen was generated with the companion repository Edendale_GXP, https://github.com/Ahmad-Mahmoudi-coder/Edendale_GXP.