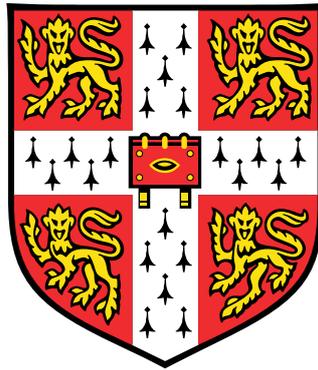


# Practical Optimisation of District Energy Systems

Representation of Technology Characteristics,  
Demand Uncertainty, and System Robustness



**Brynmor Caradog Pickering**

Department of Engineering  
University of Cambridge

This dissertation is submitted for the degree of  
*Doctor of Philosophy*



# Declaration

Brynmor Caradog Pickering  
January 2019

I hereby declare that this dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text.

Signed: \_\_\_\_\_

Date: \_\_\_\_\_

I confirm that this dissertation does not exceed the 65,000 word limit, nor the 150 figure limit, inclusive of footnotes, bibliography and appendices.



# Abstract

District energy systems are an alternative to conventional national-scale networks to meet demand in urban areas. Decentralising electricity production reduces distribution losses, while centralising thermal energy production capitalises on economies of scale. Furthermore, geographic proximity facilitates the use of waste heat from electricity production to meet thermal demand. However, it is difficult to assess which configuration of technologies will best suit a cluster of users. Mathematical optimisation techniques have been extensively researched as a method to resolve this, as they can simplify the design of investment portfolios and operation schedules for a given set of geolocated demands. However, they are not yet practically applicable.

This thesis uses the open-source, mixed integer linear programming framework Calliope to present three methodological enhancements which address model simplification, parameter uncertainty, and conflicting decision-maker objectives. These enhancements enable the practical design of district energy systems through data-centric workflows which can readily represent real system complexities in a tractable optimisation model.

This thesis first examines the impact of parameter simplification in a linear model. Piecewise linearisation is applied to nonlinear part-load energy consumption curves and a pre-processing step is developed to optimise breakpoint positioning along a piecewise curve. Second, a three-step method is proposed to handle demand uncertainty in linear models, using historical demand data. These steps are scenario generation, scenario reduction, and scenario optimisation. Using out-of-sample tests, robustness of optimal investments to unmet demand is quantified. Furthermore, system resilience to unexpected events is explored by introducing interruptions to the national electricity grid availability for a district in Bangalore, India. Scenario optimisation is extended to account for these interruptions as well as to mitigate unfavourably high levels of CO<sub>2</sub> emissions in system design. Finally, this thesis identifies eight possible decision makers, who each hold a different objective in the design of a district energy system. Optimal technology configuration and out-of-sample test results are compared across all decision-maker objectives.

These methodological enhancements demonstrate the capability of optimisation models to be reflective of reality whilst being transparent concerning the impact of simplifications, uncertainty, and conflicting objectives. Calliope is extended in this thesis to be practically applicable for district energy systems. Moreover, its extensibility facilitates the continuation of development, including possible future work into data validation and spatial dimension simplification.



## Acknowledgements

As much as a PhD is a personal project, it would not have been possible to complete this thesis without the support of so many. First and foremost I wish to acknowledge my supervisor, Ruchi, who has guided me through the mysterious world of research since my masters project in 2015, including facilitating much of my international collaboration. Advice and motivation were always forthcoming; I could not have asked for more.

I am also thankful to all those who hosted, and worked with me on my various international research trips: Shintaro and Ryoza in Tokyo, Amir and Kenrick in Bangalore, and Stefan in Zürich. More than a host, Stefan has been a close collaborator for two and a half years, aiding immeasurably in my Python programming skills, understanding of submitting jobs to supercomputers, and git workflow technique.

Data was my life source throughout this PhD and was kindly provided by the Indian Institute of Human Settlements (IIHS), the Cambridge University Living Laboratory for Sustainability, Aecom, and the Ooka laboratory at the University of Tokyo.

As well as those who helped my PhD take shape, I am grateful to the many who have ensured my sanity over the past four years: my fellow FIBEs, for our lunchtime chats and help in navigating the weird world of undertaking a PhD; my housemates and friends at 39 Mill Road, who I looked forward to spending time with at the end of each day; my family, for not pushing me to explain my research, but still chipping in on proofreading; Ewan, for believing in my work to such an extent that teams in Arup have started using Calliope; and, most importantly, Elia, the ultimate provider of treats, access to dog therapy, thorough and enlightening proofreading, and comfort throughout the highs and lows of the journey.

This research was supported by the Engineering and Physical Sciences Research Council (grant EP/L016095/1). This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (<http://www.csd3.cam.ac.uk/>), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council ([www.dirac.ac.uk](http://www.dirac.ac.uk)).

# Contents

List of Figures	ix
List of Tables	xiv
Acronyms	xvii
<b>1 Introduction</b>	<b>1</b>
1.1 Thesis overview . . . . .	3
<b>2 Literature Review</b>	<b>7</b>
2.1 Representing district energy systems . . . . .	8
2.2 Shortcomings of existing models . . . . .	25
2.3 Challenges addressed in this thesis . . . . .	32
<b>3 Energy system optimisation</b>	<b>35</b>
3.1 Mixed integer linear optimisation . . . . .	36
3.2 Calliope: an open source modelling framework . . . . .	40
3.3 Open source district energy optimisation . . . . .	57
<b>4 Exploring the linearisation of ‘reality’</b>	<b>59</b>
4.1 Linearising part-load curves . . . . .	60
4.2 Model configuration . . . . .	65
4.3 Results . . . . .	73
4.4 Discussion . . . . .	83
4.5 Conclusions . . . . .	85
<b>5 Decision making under uncertainty</b>	<b>87</b>
5.1 Handling uncertain demand . . . . .	91
5.2 Model configuration . . . . .	99
5.3 Results . . . . .	106
5.4 <b>Out-of-sample</b> testing . . . . .	125
5.5 Impact of power interruptions . . . . .	131
5.6 Discussion . . . . .	135
5.7 Conclusions . . . . .	137
<b>6 Whose objective is it anyway?</b>	<b>139</b>
6.1 Reformulating the objective function . . . . .	141
6.2 Out-of-sample tests . . . . .	166
6.3 Discussion . . . . .	170
6.4 Conclusions . . . . .	173
<b>7 Conclusions</b>	<b>175</b>
7.1 Limitations and further work . . . . .	178
<b>References</b>	<b>181</b>
<b>Appendix A State-of-the-art district energy system MILP optimisation models</b>	<b>195</b>
<b>Appendix B Case studies</b>	<b>203</b>

# List of Figures

2.1	UK electricity demand profile for a representative winter day. . . . .	11
2.2	Example of using inter-cluster storage when describing a full timeseries with <a href="#">typical days</a> . . . . .	14
2.3	Graphical representation of conventional energy technologies, as depicted in linear energy system models. . . . .	16
2.4	Graphical representation of multi-energy technologies, as depicted in linear energy system models. . . . .	19
3.1	Workflow connecting the Calliope model framework and the linear optimiser.	36
3.2	Bounded convex hull describing a two-dimensional <a href="#">linear programming</a> problem. . . . .	38
3.3	Graphical representation of an $N$ -dimensional <a href="#">linear programming</a> problem being solved by the simplex algorithm to maximise the objective function. . .	38
3.4	Illustrative two node district, with eight distinct technologies and four energy carriers. . . . .	43
3.5	Calliope supply technology. . . . .	45
3.6	Calliope conversion technology. . . . .	46
3.7	Calliope multi-energy conversion technology. . . . .	47
3.8	Representation of various multi-energy conversion technologies in Calliope. .	47
3.9	Calliope storage technology. . . . .	48
3.10	Calliope transmission technology. . . . .	49
3.11	Example of Calliope YAML model formulation in which a <a href="#">combined heat and power plant</a> technology is defined. . . . .	53
3.12	Internal Calliope workflow, including third party library dependencies. . . .	54
3.13	Figures produced using Calliope plotting functions. . . . .	56
4.1	Example of nonlinear curves. . . . .	61
4.2	Application of <a href="#">convex bounding sets</a> to describe a nonlinear curve. . . . .	63
4.3	Graphical representation of piecewise linearisation using <a href="#">special order set constraint of type 2</a> . . . . .	64
4.4	Nonlinear characteristic curves for various energy technologies. . . . .	67
4.5	Energy demand and fuel pricing for case study 1. . . . .	68
4.6	Schematic of supply and storage technologies for meeting multi-energy demand in case study 1. . . . .	69
4.7	Graphical representation of case study 2 district network. . . . .	71
4.8	Progression of curve linearisation used in tests, from nonlinear to straight line.	72

---

4.9	Comparison of optimal operation schedules for meeting hot water demand in case study 1. . . . .	76
4.10	Comparison of optimal capacity and dispatch between different part-load characteristic linearisation methods in case study 2. . . . .	77
4.11	Impact of optimised and equidistant piecewise curves on ex-ante <b>linearisation error (LE)</b> of part-load curves of technologies considered in case study 2. . . . .	78
4.12	Comparison of optimal capacity and dispatch as the number of piecewise linear breakpoints is increased from two (rated efficiency) to six, for case study 2 operating in summer and winter. . . . .	82
5.1	Comparison of parametric and nonparametric sampling methods as a means to produce daily profiles for demand. . . . .	92
5.2	Example of <b>kernel density estimation</b> , applied to a single feature observation set. . . . .	93
5.3	Representation of the <b>conditional value at risk</b> risk measure. . . . .	97
5.4	Comparison of quantity of unmet demand incurred in various scenarios for penalty rates ranging from 100 to 10 <sup>6</sup> <b>Indian rupees/kWh</b> , applied to a test model of the Bangalore case study district. . . . .	99
5.5	Case study districts. . . . .	100
5.6	Bangalore case study input demand profiles, grouped by sampling cluster and energy type. . . . .	101
5.7	Cambridge case study input demand profiles, grouped by sampling cluster, archetype, and energy type. . . . .	104
5.8	Total hourly district demand in one typical year, as sampled for 500 scenarios. . . . .	108
5.9	<b>Typical day</b> demand profiles for both Bangalore and Cambridge case study districts. . . . .	109
5.10	Comparison of annual energy demand distributions between that derived from the full timeseries and from the clustered timeseries of 500 generated scenarios in Cambridge and Bangalore. . . . .	111
5.11	Optimal technology energy capacity for the Bangalore case study scenario 70 model, following increasingly aggressive timeseries aggregation. . . . .	112
5.12	Total district demand compared to independently optimal <b>objective function value</b> for 500 demand scenarios. . . . .	114
5.13	Installed capacity of technologies to achieve the optimal <b>objective function value</b> in both mean and <b>scenario optimisation</b> cases. . . . .	117
5.14	Installed capacity and capacity factor for conventional technologies chosen to achieve the optimal <b>objective function value</b> in the mean, 10-scenario <b>scenario optimisation</b> , and 500 independent scenario model runs for both Cambridge and Bangalore case studies. . . . .	118
5.15	Contribution of investment and operation costs to the <b>objective function value</b> of the mean model, 10-scenario <b>scenario optimisation</b> model, and all 500 independent scenario models. . . . .	119

5.16	Distribution of total system cost resulting from running models using risk-neutral <a href="#">scenario optimisation</a> , risk-averse <a href="#">scenario optimisation</a> , and risk-unaware optimisation on all 500 independent scenarios. . . . .	121
5.17	Optimal energy capacity resulting from running the mean model compared to using risk-neutral <a href="#">scenario optimisation</a> and risk-averse <a href="#">scenario optimisation</a> . . . . .	122
5.18	Result of sensitivity analysis on the levels of risk aversion $\beta$ and number of reduced scenarios used in risk-averse <a href="#">scenario optimisation</a> , applied to the Bangalore case study. . . . .	123
5.19	Depiction of a rolling horizon optimisation. . . . .	126
5.20	System unmet thermal and electricity demand distributions when running <a href="#">out-of-sample</a> optimisation tests on optimal technology portfolios resulting from single scenario optimisations and risk-neutral <a href="#">scenario optimisation</a> in both Cambridge and Bangalore case studies. . . . .	128
5.21	System unmet thermal and electricity demand distributions when running <a href="#">out-of-sample</a> optimisation tests on risk-neutral <a href="#">scenario optimisation</a> and risk-averse <a href="#">scenario optimisation</a> optimal technology portfolios in both Cambridge and Bangalore case studies. . . . .	129
5.22	System unmet cooling and electricity demand distributions when running <a href="#">out-of-sample</a> optimisation tests for the optimal technology portfolios resulting from the same demand scenario optimised with either no timeseries aggregation or aggregation to 12 <a href="#">typical days</a> . . . . .	129
5.23	Results of sensitivity analysis on the unmet cooling and electricity demand realised in <a href="#">out-of-sample</a> optimisation tests for varying levels of risk aversion $\beta$ and number of reduced scenarios, applied to the Bangalore case study. . . . .	130
5.24	Distribution of sampled power interruptions across the Bangalore case study year, based on historical unscheduled interruption data in Bangalore. . . . .	132
5.25	System unmet cooling and electricity demand distributions when running <a href="#">out-of-sample</a> optimisation tests which include electricity power supply interruptions, applied to the Bangalore case study models. . . . .	132
5.26	Comparison of mean model a) installed capacity and b) annual electricity production to results given by various models aimed at improving power interruption resilience. . . . .	134
5.27	System unmet cooling and electricity demand distributions when running <a href="#">OOS</a> optimisation tests, including the introduction of electricity power supply interruptions, on Bangalore case study models which have incorporated power interruption resilience. . . . .	135
6.1	Representation of a convex hull, containing non-dominant solutions resulting from multi-objective optimisation, and a Pareto front on which the dominant solutions lie. . . . .	143

6.2	Comparison of optimal investment and operation costs for Cambridge and Bangalore case studies when balancing operation and investment costs , minimising only investment cost, and minimising only operation cost in the objective function. . . . .	145
6.3	Investment-operation cost Pareto front following monetary cost minimisation of both case studies with $\epsilon$ -constrained investment cost. . . . .	147
6.4	Installed energy capacity following monetary cost minimisation of both case studies with $\epsilon$ -constrained investment cost. . . . .	147
6.5	Installed capacity of technologies to achieve the optimal <b>objective function value (OFV)</b> in both cost minimisation and carbon emission minimisation models.	149
6.6	Timeseries energy production and consumption in the Bangalore case study for technologies operating in each typical day, as given by the optimal model solution following carbon emission minimisation. . . . .	150
6.7	Timeseries energy production and consumption in the Cambridge case study for technologies operating in each typical day, as given by the optimal model solution following carbon emission minimisation. . . . .	150
6.8	Cost-carbon Pareto front following monetary cost minimisation of both case studies with $\epsilon$ -constrained carbon emissions. . . . .	154
6.9	Installed energy capacity following monetary cost minimisation of both case studies with $\epsilon$ -constrained carbon emissions. . . . .	155
6.10	Investment cost-carbon Pareto front following carbon emission minimisation of both case studies with $\epsilon$ -constrained investment cost. . . . .	155
6.11	Installed energy capacity following carbon emission minimisation of both case studies with $\epsilon$ -constrained investment cost. . . . .	156
6.12	Installed energy capacity following monetary cost minimisation of both case studies with a percentage limit set on the sum of district thermal demand which can be met by building-level conventional technologies. . . . .	158
6.13	Cost minimised vs. carbon minimised <b>objective function value</b> for 500 generated scenarios in the Bangalore case study. . . . .	162
6.14	Comparison of optimal technology capacity and annual energy production for Bangalore case study baseline and <b>scenario optimisation</b> models. . . . .	164
6.15	Distribution of total system monetary cost and carbon emissions resulting from the minimisation of cost (including the cost of carbon), using risk-neutral <b>scenario optimisation</b> and risk-averse <b>scenario optimisation</b> Bangalore case study models. . . . .	165
6.16	Distribution of unmet cooling and electricity demand realised when running <b>out-of-sample</b> tests on the result of cost minimisation with $\epsilon$ -constrained investment cost in the Bangalore case study district. . . . .	167
6.17	Distribution of unmet cooling and electricity demand realised when running <b>out-of-sample</b> tests on the result of carbon minimisation with $\epsilon$ -constrained investment cost in the Bangalore case study district. . . . .	167

---

6.18	Distribution of unmet cooling and electricity demand realised when running <b>out-of-sample</b> tests on the result of cost minimisation with decreasing allowed share of cooling demand met by <b>electric chillers</b> in the Bangalore case study district. . . . .	168
6.19	Distribution of unmet cooling and electricity demand realised when running <b>out-of-sample</b> tests on the result of cost minimisation with $\epsilon$ -constrained carbon emissions in the Bangalore case study district. . . . .	168
6.20	Distribution of unmet cooling and electricity demand realised when running <b>out-of-sample</b> tests on the result of baseline, single scenario optimisation and risk-neutral, 10-scenario <b>scenario optimisation</b> in the Bangalore case study district. . . . .	169
6.21	Distribution of unmet cooling and electricity demand realised when running <b>out-of-sample</b> tests on the result of 10-scenario risk-neutral <b>scenario optimisation</b> and risk-averse <b>scenario optimisation</b> in the Bangalore case study district. . . . .	170
B.1	Case study 1 energy demand and fuel pricing. . . . .	203
B.2	Case study 1 technology connection schematic. . . . .	204
B.3	Case study 2 district network. . . . .	206
B.4	Case study 2 input demand. . . . .	207
B.5	Case study 3 district network. . . . .	211
B.6	Typical day demand profiles for case study 3. . . . .	212
B.7	Case study 4 district network. . . . .	215
B.8	Typical day demand profiles for case study 4. . . . .	216

# List of Tables

1.1	Overview of thesis case studies. . . . .	5
3.1	<a href="#">Mixed integer linear programming</a> model solution time when run using different linear optimisers. . . . .	40
3.2	Decision variables defined in Calliope. . . . .	42
3.3	Sets and subsets defined in Calliope. . . . .	43
3.4	Valid parameters for constraining a supply technology. . . . .	44
3.5	Valid parameters for constraining a conversion technology. . . . .	45
3.6	Valid parameters for constraining a multi-energy conversion technology. . . . .	46
3.7	Valid parameters for constraining a storage technology. . . . .	48
3.8	Valid parameters for constraining a transmission technology. . . . .	49
3.9	Valid parameters for constraining a demand technology. . . . .	50
3.10	Valid parameters for applying costs to decision variables in Calliope. . . . .	51
4.1	Supply technologies and their consumption/production energy. . . . .	68
4.2	Case study 2 building characteristics. . . . .	70
4.3	Capacity costs of supply technologies in case study 2. . . . .	71
4.4	Run configuration options for comparing nonlinear curves to their <a href="#">single line efficiency</a> and piecewise linearised counterparts. . . . .	72
4.5	<a href="#">Objective function value</a> and optimisation solution time for case study 1, comparing linear, piecewise linear, and nonlinear technology representation. . . . .	74
4.6	<a href="#">Objective function value</a> and optimisation solution time for case study 2, comparing linear and piecewise linear technology representation. . . . .	74
4.7	<a href="#">Objective function value</a> for all breakpoint comparison run configurations. . . . .	80
4.8	Model runtime for all configurations, including pre-processing and subsequent optimisation. . . . .	80
4.9	Comparison between piecewise and rated efficiency results of distribution network between, and storage capacity at, demand locations of case study 2. . . . .	81
5.1	Bangalore district node details. . . . .	100
5.2	Cambridge district node details. . . . .	102
5.3	Dates corresponding to Cambridge term times and vacations, as used to define sampling clusters for the Cambridge case study. . . . .	103
5.4	Calculated bandwidths for each <a href="#">kernel density estimation</a> subset in the Cambridge case study input dataset. . . . .	106

5.5	Comparison of solution time, total system costs, and annual energy demand associated with the optimal solution of the Bangalore case study scenario 70 model, following increasingly aggressive timeseries aggregation. . . . .	112
5.6	Description of 10 reduced scenarios chosen to represent the full set of 500 generated scenarios. . . . .	116
5.7	Investment cost for optimal solutions derived from Bangalore models aimed at improving power interruption resilience. . . . .	135
6.2	Cost of carbon range, as given by various sources. . . . .	152
6.3	Optimal investment costs, operation costs, and carbon emissions for baseline carbon and cost models in both Cambridge and Bangalore case studies. . . . .	153
6.4	Monetary cost and carbon emissions following monetary cost minimisation of both case studies with a percentage limit set on the sum of district thermal demand which can be met by building-level conventional technologies. . . . .	158
6.5	Percentage share of building thermal demand met by building-level conventional technologies in each node of both case study districts. . . . .	159
A.1	Model reference table. . . . .	195
A.2	Glossary of columnwise abbreviations used in Table A.3. . . . .	197
A.3	Details of models studied in reviewed literature. . . . .	198
B.1	Case study 1 technology characteristics. . . . .	204
B.2	Technologies associated with each decision variable in case study 1. . . . .	205
B.3	<b>MILP</b> problem size for case study 1, as reported by IBM ILOG CPLEX. . . . .	205
B.4	Case study 2 district node details. . . . .	206
B.5	Case study 2 technology characteristics. . . . .	208
B.6	Details of distribution technologies in case study 2. . . . .	208
B.8	Technologies associated with each decision variable in case study 2. . . . .	210
B.9	<b>MILP</b> problem size for case study 2, as reported by IBM ILOG CPLEX. . . . .	210
B.10	Case study 3 district node details. . . . .	211
B.11	Case study 3 technology characteristics. . . . .	213
B.12	Technologies associated with each decision variable in case study 3. . . . .	214
B.13	<b>MILP</b> problem size for case study 3, as reported by IBM ILOG CPLEX. . . . .	214
B.14	Case study 4 district node details. . . . .	216
B.15	Case study 4 technology characteristics. . . . .	217
B.16	Technologies associated with each decision variable in case study 4. . . . .	218
B.17	<b>MILP</b> problem size for case study 4, as reported by IBM ILOG CPLEX. . . . .	219



# Acronyms

$\epsilon$ DE	epsilon constraint differential evolution.
$\mu$ CHP	micro <a href="#">CHP</a> .
AHP	air source heat pump.
AR	absorption refrigerator.
B-CCHP	biomass fuelled <a href="#">CCHP</a> .
CBS	convex bounding sets.
CCHP	combined cooling, heat and power plant.
CHP	combined heat and power plant.
COP	coefficient of performance.
CVaR	conditional value at risk.
D-CCHP	diesel fuelled <a href="#">CCHP</a> .
DG	diesel generator.
EC-ECh	energy centre <a href="#">ECh</a> .
ECh	electric chiller.
FC	fuel cell.
fPCA	functional principal component analysis.
GBP	British pound sterling.
GridE	national grid electricity.
GridNG	national grid natural gas.
GSHP	ground source heat pump.
GT	gas turbine.
HP	heat pump.

<b>HRAR</b>	heat recovery absorption refrigerator.
<b>HTP</b>	heat-to-power ratio.
<b>ICE</b>	internal combustion engine.
<b>INR</b>	Indian rupees.
<b>JPY</b>	Japanese yen.
<b>KDE</b>	kernel density estimation.
<b>LE</b>	linearisation error.
<b>LP</b>	linear programming.
<b>MILP</b>	mixed integer linear programming.
<b>NGB</b>	natural gas boiler.
<b>OFV</b>	objective function value.
<b>OOS</b>	out-of-sample.
<b>PDF</b>	probability density function.
<b>PV</b>	solar photovoltaic panel.
<b>RMSE</b>	root-mean-square error.
<b>RO</b>	robust optimisation.
<b>SE</b>	Stirling engine.
<b>SG</b>	scenario generation.
<b>SLE</b>	single line efficiency.
<b>SLSQP</b>	sequential least squares programming.
<b>SO</b>	scenario optimisation.
<b>SOS2</b>	special order set constraint of type 2.
<b>SR</b>	scenario reduction.
<b>ST</b>	solar thermal panel.
<b>StoreE</b>	electrical battery storage.
<b>StoreH2</b>	hydrogen storage.
<b>StoreT</b>	thermal energy storage.
<b>StoreT-C</b>	cold water <a href="#">StoreT</a> .

<b>StoreT-H</b>	hot water <a href="#">StoreT</a> .
<b>TD</b>	typical day.
<b>TSF</b>	timeseries scaling factor.
<b>USD</b>	United States dollar.
<b>VaR</b>	value at risk.
<b>WT</b>	wind turbine.



# Chapter 1

## Introduction

There are growing calls to update how global energy demand is met (IPCC, 2014). Particularly, this pertains to the reduction in dependence on fossil fuels in favour of zero, or near zero, carbon emitting energy sources. Buildings are the largest consumers of energy, constituting approximately 75% of natural gas consumption and approximately 70% of electricity consumption in the UK (BEIS, 2018a). As such, methods to reduce demand, increase supply efficiency, and change energy sources are a longstanding research area. Currently, building energy provision is characterised by direct consumption of nationally generated electricity and locally generated heat, with the latter produced from nationally distributed natural gas. However, in response to the need for new energy systems, a paradigm shift may be necessary.

District energy systems offer a promising avenue for building energy provision. In such systems, end-use energy demand is met by generation at close geographic proximity. This allows electricity generation to become decentralised from the national grid and thermal generation to be centralised relative to building-level generation. The small distance required to distribute energy in a district reduces electricity distribution losses, compared to national-scale transmission (Pepermans *et al.*, 2005). The potential advantages of this are more pronounced in less economically developed nations, where overall energy losses caused by poor national grid infrastructure can be three times greater than recorded in Europe (IEA Statistics, 2014). Furthermore, the inadequacies of national grid infrastructure can cause intermittency in electricity supply, which will have negative economic impacts (Dollar *et al.*, 2005).

In addition to the economic benefits of decentralised electricity generation, centralised thermal energy generation can capitalise on economies of scale (Söderman, 2007) and otherwise waste energy streams (Rezaie and Rosen, 2012). In meeting greenhouse gas emission reductions in the European Union, district heating systems could offer an equally viable, if not cheaper, alternative to other methods such as the electrification of decentralised heat generation (Connolly *et al.*, 2014). Central to this economic benefit are **combined heat and power plants (CHPs)**, which can produce electricity conventionally, using a gas turbine, for instance, whilst their exhaust heat is captured and used to heat water. This water is then distributed, via a piped network, to buildings in a district, to provide space heating that would otherwise only be available from building-level technologies. Many industrialised cities have

CHP systems (United Nations Environment Programme, 2014). Indeed, districts within cities benefit most from such systems, as return on investment increases with population density (IEA, 2014).

Although the potential benefits of district energy systems are evident in the literature, the prevalence of such systems has not markedly increased in recent years. In the UK, the total capacity of CHP schemes even reduced between 2013 and 2017, from 5,924MWe to 5,835MWe (BEIS, 2018a). In contrast, research into modelling such systems has experienced a six-fold increase between 2012 and 2018<sup>1</sup>. Barriers to implementation are caused primarily by system cost. For example, electricity generating district systems are more expensive to install than their national scale counterparts and may limit the available primary fuel choice to only the most expensive (Pepermans *et al.*, 2005). The installation of thermal energy distribution networks is also costly (Omu *et al.*, 2013), exacerbated by the need to replace existing infrastructure in heavily populated areas (Rezaie and Rosen, 2012).

District energy systems inherit complexities from both building-scale systems and national-scale systems. They consider multiple, interconnected energy streams whilst also requiring the distribution of energy to consumers with varying spatial and temporal characteristics. Therefore, increasing the uptake of centralised district energy systems and realising their potential benefits requires more sophisticated tools. Mathematical optimisation is one such tool which is often associated with district energy system design (Keirstead *et al.*, 2012) and has been researched for use in district design for over three decades (Gustafsson *et al.*, 1987). Optimisation is the process of making the best of something; an optimal energy system is one which minimises a given objective whilst meeting the physical constraints of the system. Given the complex possible interactions of district energy systems, it can be difficult to understand what combination of technologies would yield an optimal solution. Hence, computational methods are required, reducing the burden on the designer to only describing key constraints of the system.

Many studies exist with a focus on designing district energy systems using optimisation techniques (Keirstead *et al.*, 2012; Sameti and Haghghat, 2017). These studies often investigate the minimisation of system monetary cost while meeting the demand of a number of buildings by either centralised or decentralised technologies. Some studies also minimise greenhouse gas emissions or energy use (Sameti and Haghghat, 2017). However, system scale, resolution, location, available technologies, and recommendations vary between studies. This variation is coupled with limited model and data transparency, making it difficult to ascertain the validity of the modelling methods used and the benefit of implementing district systems (DeCarolis *et al.*, 2012). This may explain the limited use of mathematical optimisation in industry, in favour of heuristic methods.

---

<sup>1</sup>The increase in research activity has been measured by comparing the number of publications available when searching for '(urban OR city) energy model' in topic or title using the ISI Web of Knowledge database in November 2018 to the same search carried out in 2012 (Keirstead *et al.*, 2012).

## 1.1 Thesis overview

### 1.1.1 Hypothesis, aims, and tasks

The central hypothesis of this thesis is that current district energy models fall short of being practically applicable, owing primarily to the following reasons:

- 1 *Models are not structured in a transparent manner and they cannot be validated.*
- 2 *Models do not fully recognise the simplifications made in preparing a linear representation of real systems.*
- 3 *Models do not understand the impact of deterministic decision-making on the feasibility of investment decisions.*

As such, the aims of this thesis are to:

1. Develop a transparent, extensible model;
2. Test the influence of simplifying technology part-load characteristics on the optimum;
3. Investigate techniques for improving district energy model robustness to demand uncertainty and power interruptions; and
4. Compare the impact of different objectives on model feasibility when exposed to uncertainty.

To fulfil these aims, the following tasks were developed:

1. Design and deploy an open-source piece of software, which a decision maker can use to prepare, optimise, and analyse a district energy system model;
2. Create piecewise linearised representations of technology part-load consumption curves, for direct comparison to a nonlinear model and a linear model using rated efficiency;
3. Formulate a method to handle demand uncertainty in district energy systems; and
4. Design **out-of-sample** tests to consistently measure the robustness of investment decisions to unexpected future scenarios.

### 1.1.2 Thesis structure

This thesis begins in Chapter 2 by reviewing the current state-of-the-art in district energy modelling. A district energy model is broken down to its component parts, with each part being discussed in turn. The way in which studies from recent years have represented each component is compared, to better understand the breadth of the field of linear district energy system optimisation models. Shortcomings of existing studies are discussed in three parts:

model simplifications, model determinacy, and model transparency. These shortcomings inform the hypothesis on which this thesis is based. The studies pushing frontiers to overcome these shortcomings are discussed alongside those which are examples of the current paradigm. Finally, the focus of this thesis is discussed, directly linking the thesis hypothesis, aims, and tasks.

[Mixed integer linear programming \(MILP\)](#) and the required mathematical formulation to construct an energy system optimisation model is presented in [Chapter 3](#). In doing so, the open-source modelling framework *Calliope* is detailed. *Calliope* has been extended in this thesis to provide the functionality required to meet the thesis aims and to expand its capabilities as a practical tool for decision making. This includes the automation of aggregation tasks, improved clarity in defining an energy system, shareable data packaging, and interactive data visualisation.

[Chapter 4](#) details an exploratory analysis of piecewise linearisation. Namely, the piecewise linearisation technique is applied to nonlinear technology part-load consumption curves. Two parallel models are developed by which piecewise linearisation can be compared to the fully nonlinear representation of the curves. To improve on piecewise linear model tractability, the linear model is expanded to capitalise on the strictly convex upward or downward nature of the curves under study and to implement a pre-processing step which can be used to automatically generate piecewise curves using the fewest possible breakpoints along the curves.

A three-step method by which uncertainty can be accounted for in district energy system models is detailed in [Chapter 5](#). The method encompasses data-driven, multi-dimensional scenario generation; scenario reduction using independently optimal objective function values; and tractable [scenario optimisation \(SO\)](#). Historical energy demand data from an office space in Bangalore, India, and commercial/academic spaces in Cambridge, UK are used to inform scenario generation. Optimal technology capacity is sought using varying levels of demand uncertainty awareness. These technology portfolios are then directly compared using [out-of-sample \(OOS\)](#) tests, which aim to quantify robustness. Portfolios which can ensure demand is met, even when exposed to new demand data, are considered robust. Furthermore, the historical prevalence of unexpected interruptions to national grid power in Bangalore is quantified. The almost daily occurrence of these interruptions is used to extend both [SO](#) and [OOS](#) tests to quantify and improve system resilience.

The impact of objective function choice is examined in [Chapter 6](#), using the case studies introduced in [Chapter 5](#). Technology investment is compared when varying the relative importance of investment and operation costs, carbon emissions and monetary costs, and dependence on a district thermal network. The inadvertent impact of these investment portfolios is quantified by exposing them to new demand scenarios.

Finally, the methods of [chapters 5 and 6](#) are combined by the consideration of carbon emissions as a risk factor in [SO](#). The aim of this combination is to converge on technology investment portfolios which are cost effective whilst robust to demand uncertainty and the realisation of excessive carbon emissions. [Chapter 7](#) concludes by drawing out the

contributions of this thesis to the field of linear district energy optimisation, with particular emphasis on achieving the thesis aims.

### 1.1.3 Case studies

Four case studies are used in this thesis, each with different properties and uses. Table 1.1 provides an overview of each case study, including their key properties and the chapter(s) in which they are used. Greater detail on each case study can be found in the relevant chapters, as well as in Appendix B.

Case study 1 is represents a hotel in Japan, optimised over 24 hours to determine only technology operation, not investment. It has hot water, cooling, and electricity demand.

Case study 2 is an illustrative district in the UK, optimised over four typical days to determine both investment and technology operation. It takes many of the technology characteristics from case study 1, but with costs and geographic parameters updated for the new context. It represents domestic and commercial properties, with demand for electricity, heat, and cooling. A power plant is also proposed on the network, which can distribute heat, gas, and electricity to the demand nodes.

Case study 3 is an illustrative district in Bangalore, India, based on a commercial area of the city, optimised over 12 typical days to determine both investment and technology operation. It represents a district cooling network amongst office buildings only, with the possibility of an energy centre to produce cold water at a large scale. New technology characteristics, including costs, are used in this case study.

Case study 4 is based on the masterplan of a proposed district in Cambridge, UK, optimised over six typical days to determine both investment and technology operation. It comprises laboratories and offices with heat and electricity demand and is the largest district network. The viability of a proposed energy centre and district heating network is analysed in this case study. Technology characteristics draw from those in case study 2.

Table 1.1 Overview of thesis case studies. Abbreviations are given on pages xv and xvi.

	Spatial size	Temporal size	Location	Number of technologies	Thesis chapter
Case study	1 1 building	1 day	Japan	9 supply 2 storage 0 distribution	4
	2 10 buildings, aggregated to 2 demand nodes	1 year, aggregated to 4 typical days	UK	10 supply 3 storage 3 distribution	4
	3 17 buildings, aggregated to 11 demand nodes	1 year, aggregated to 12 typical days	Bangalore, India	6 supply 2 storage 2 distribution	5, 6
	4 46 buildings, aggregated to 39 demand nodes	1 year, aggregated to 6 typical days	Cambridge, UK	5 supply 2 storage 3 distribution	5, 6



# Chapter 2

## Literature Review

This chapter reviews recent and state-of-the-art in district energy optimisation models. In section 2.1, a detailed analysis is undertaken of the following model components required to define a district: 1) system resolution, 2) representation of supply technologies, and 3) multi-energy demand. The review focusses on the need for a model to represent a real system, as well as the steps taken within the existing literature to address the difficulty of doing so. In Section 2.2, model simplification, determinacy and lack of transparency are discussed as key shortcomings of the current paradigm in district energy system optimisation. Finally, in Section 2.3 the focus of this thesis is placed within the context of the current paradigm and its shortcomings.

Collating studies on district energy systems requires broadening the terminology used to define such systems. A district energy system is defined in this thesis as one in which useful energy is generated in close geographical proximity to demand nodes, allowing synergies between multiple energy carriers to be exploited. These energy carriers are predominantly electrical and thermal (high and low temperature). District systems are ultimately connected to national energy networks and can thus be considered subsystems within a complete national energy system. For this reason, recent reviews of the existing literature have defined systems within their scope as 'distributed' (Alarcon-Rodriguez *et al.*, 2010; Chicco and Mancarella, 2009), 'decentralised' (Hiremath *et al.*, 2007), or 'microgrid' (Gu *et al.*, 2014). However, the terms 'distributed' and 'decentralised' can cause confusion. An energy centre within a district system is a centralised system when compared to building-level energy provision, but is decentralised compared to the national grid infrastructure. Although the term 'microgrid' is perhaps less confusing in placing district energy systems within their spatial context, this definition may also suggest that they are purely electrical systems (Marnay *et al.*, 2015). Instead, some reviews have used the terms 'community' and 'urban' to define energy systems, bounding the models geographically. A community is considered to be a system smaller than 10km<sup>2</sup> (Huang *et al.*, 2015) while urban energy systems include both 'district' and entire city systems (Keirstead *et al.*, 2012; Reinhart and Cerezo Davila, 2016). Although a district system could reasonably be the size of a small city, the loose terminology of 'urban' excessively broadens the available literature. Mancarella (2014) ignores geographic

bounding entirely by using the definition ‘multi-energy systems’. But, in doing so, their review encompass both building-level studies and regional models which connect districts.

Recent reviews of optimisation models have preferred to use the definition of *district* energy systems (Allegrini *et al.*, 2015; Olsthoorn *et al.*, 2016; Sameti and Haghighat, 2017). Of these, Sameti and Haghighat (2017) offer the most comprehensive review of existing district energy system optimisation models, including both linear and nonlinear systems within their study.

The scope of this review is limited to studies which optimise district multi-energy systems using linear programming algorithms. Within this scope, a core set of 60 models has been identified, of which 46 have been published in the last four years and can be considered state-of-the-art in the field. Table A.3 in Appendix A provides the reader with a summary for each model, which collates the components chosen based on their relevance to this review.

## 2.1 Representing district energy systems

### 2.1.1 System size

Translating a physical system to a digital representation requires discretely defining much of what is continuous in reality. However, discretisation leads to loss of fidelity. This phenomenon can be likened to digital displays where highly pixelated (low resolution) images are seen to be a poor reflection of reality. Increasing resolution improves the image, but at the cost of increased power consumption to compute each pixel’s colour and to power more light emitting diodes. Worse still, a larger display leads to a polynomial increase in power consumption, as the number of pixels must increase in both dimensions to maintain the requisite resolution. Much like digital displays, there is a desire for high resolution, large district energy system models. Although system resolution has improved, the computational expense is ever present. There are two dimensions by which district systems can be discretised: in space and in time. In the literature, there is no agreed resolution or size, with most studies choosing to compromise on one of the dimensions.

#### The spatial dimension

Districts are composed of distinct buildings, to which energy is distributed. Studies which model buildings individually are necessarily small. Omu *et al.* (2013) and Casisi *et al.* (2009) only considered six buildings in a district smaller than 0.02km<sup>2</sup>. Indeed, fewer than 20 buildings is common when they are modelled independently (Casisi *et al.*, 2009; Li *et al.*, 2016a; Maréchal *et al.*, 2008; Mehleri *et al.*, 2012; Morvaj *et al.*, 2016; Voll *et al.*, 2013). There are models that exceed this, with Söderman (2007) most notably incorporating up to 53 buildings in a district. However, most of these larger districts are optimised only for energy system operation (Bucciarelli *et al.*, 2018; Cesena and Mancarella, 2018; Good and Mancarella, 2017; Orehounig *et al.*, 2015; Reddy, 2017). In such models, no decisions need to be made about the technology portfolio in which to invest. As a compromise, Orehounig *et al.* (2015) considered a range of scenarios consisting of different technology portfolios and

independently optimised the operation schedule for a 29 building district for each scenario. Where optimisation includes technology capacity selection, only Voll *et al.* (2015) maintain a full hourly resolution over a year (8760 timesteps) when modelling 39 consumers in a district. Other models reduce the time resolution to 20 or fewer timesteps (Haikarainen *et al.*, 2014; Khir and Haouari, 2015; Söderman, 2007; Tanaka *et al.*, 2017; Weber and Shah, 2011).

To increase the number of buildings in the district while including technology capacity optimisation, it is most common to reduce the spatial resolution of a model. In the first instance, buildings are aggregated into nodes which incorporate a small number of adjacent buildings. For instance, when modelling an industrial site, Buoro *et al.* (2013) modelled nine energy ‘users’ to integrate at least 20 buildings into a 5km district network. Beyond this level of aggregation, it is likely that the district network would be drastically simplified. Aggregating 75 buildings into five nodes (Omu *et al.*, 2015) or 137 buildings into seven nodes (Ameri and Besharati, 2016) does not allow a model to accurately capture any variability of demand across individual buildings. Jennings *et al.* (2014) simplified even further by aggregating a London borough of 92,170 buildings into 19 interconnected nodes while other studies have modelled entire cities with 12 (Söderman and Pettersson, 2006) or 26 (Haikarainen *et al.*, 2014) ‘consumers’.

At its most extreme, aggregation of buildings into nodes leads to a lack of any network consideration in a model. Energy balance is still maintained, but with the assumption that there is no distance between energy generation and consumption. Garcia and Weisser (2006) modelled the entirety of Grenada as a single demand node when optimising the operation of a wind-diesel-storage electrical system. Heat networks are also aggregated to single nodes, both in modelling city-scale systems (Christidis *et al.*, 2012; Farzaneh *et al.*, 2016; Fazlollahi *et al.*, 2014) and areas of cities at the borough level (Gabrielli *et al.*, 2018; Ren *et al.*, 2010). Even when the size of the district would be sufficiently small to include a network, consisting of only four or five buildings, some studies choose to model only a single node (Koltsaklis *et al.*, 2014; Kotzur *et al.*, 2018a; Kotzur *et al.*, 2018b; Wouters *et al.*, 2014). At this spatial resolution, a district energy model is indistinguishable from any other spatial scale. There are some national-level and building-level energy systems which are almost interchangeable, particularly as the latter also models multi-energy demand.

For centralised energy to be provided within a district, an energy centre must be sited and energy distribution must be physically placed into the system. Consumers or producers are point nodes in some studies (Haikarainen *et al.*, 2014; Li *et al.*, 2016a; Morvaj *et al.*, 2016; Söderman, 2007; Voll *et al.*, 2013), but a lack of understanding around the physical footprint of nodes can ignore potential limitations. For instance, reducing the footprint of a necessarily large district-scale storage tank would require a taller tank, increasing the height to diameter (H/D) ratio. While a larger H/D ratio does improve utilisation of tank volume, it increases tank heat loss (Nielsen, 2005). Where existing infrastructure is modelled, space must be found within the boundary of the district. Omu *et al.* (2015) pinpointed existing building-level combined cooling, heat and power plants (CCHPs) on a site, and designated an area of sparse existing development for a proposed district combined heat and power plant (CHP). Similarly, Buoro *et al.* (2014) used undeveloped space of an industrial district to site a CHP.

New developments have greater freedom to fit a district system optimally around demand nodes. For instance, Keirstead *et al.* (2010) considered a proposed eco-town using their model *SynCity*. By including building occupant capacity and location endogenously in the model, they were able to find the best configuration of buildings, making best use of district thermal energy. This same eco-town has also been used in other studies, but with fixed physical infrastructure reduced to point nodes (Jennings *et al.*, 2014; Voll *et al.*, 2015; Weber and Shah, 2011).

Not only do energy technologies need physical space in which to be sited, but distribution networks must be routed. In single node models, such as the 500,000 buildings modelled by Fazlollahi *et al.* (2014), routing thermal energy to all consumers is assumed possible. In studies which model the district network, it is common to use the road-following approach, in which the distribution network is assumed to follow the existing road network (Mavromatidis *et al.*, 2018a). Again, by modelling a new development, Keirstead *et al.* (2010) did not have to take the road network into account as the district was undeveloped. However, Weber and Shah (2011) noted that road-following restricts the interconnection of nodes in an existing district. In fact, Voll *et al.* (2013) did not extend a district cooling network to one building in their networked model as a major existing road would prevent the pipe from being laid. Morvaj *et al.* (2016) examined the impact that road-following has on the optimal solution, using an abstract network of twelve buildings. When not constrained to road-following, any building could connect to any other building in the network. Morvaj *et al.* (2016) found that if road-following was required, buildings were less likely to connect into a centralised network. Thus, models which do not consider the routing of distribution networks may overstate the viability of a centralised system (e.g. Akbari *et al.*, 2016; Bucciarelli *et al.*, 2018; Li *et al.*, 2016a; Mehleri *et al.*, 2012; Reddy, 2017; Wouters *et al.*, 2015).

### **The temporal dimension**

An energy system model is described along several dimensions, including the number of nodes, technologies, energy carriers, and steps in time. The longest of these dimensions is, almost universally, time. While the a district rarely increases above 20 nodes, the time horizon for optimisation is rarely below 20 time steps. The preferred resolution is hourly, spanning over a year for planning models and over a day for operation scheduling models. But, in planning models this equates to 8760 elements in time dimension. Only a few models manage to maintain this resolution. Orehounig *et al.* (2015) and Capuder and Mancarella (2014) do so by heuristically selecting technology portfolios, and avoid the complexity of including technology capacity as a decision variable. Other models reduce spatial resolution down to a single node (Gabrielli *et al.*, 2018; Garcia and Weisser, 2006; Kotzur *et al.*, 2018a; Kotzur *et al.*, 2018b). Only five studies in the scope of this review included an entire district network, full temporal resolution for a full year, and endogenous technology capacity in their models (Buoro *et al.*, 2014; Keirstead *et al.*, 2010; Mavromatidis *et al.*, 2018a; Voll *et al.*, 2015). It is not surprising that so few models maintain both the temporal resolution and scale required to reach 8760 time steps since linear programming algorithms tend towards solving in a polynomial time, relative to the problem size (Megiddo, 1986).

In optimising the operation schedule of a model, higher temporal resolution can better capture peaks and troughs in demand. However, to maintain model tractability, a lower temporal resolution may be required. Downsampling is the process of aggregating higher resolution data into a lower resolution set, usually by averaging the values of higher resolution data points, such as the average of 12 5-minute data points into a 1-hour data point. Figure 2.1 illustrates the case of downsampling from a resolution of 5 minutes to 1 or 12 hours. The one hour dataset captures the key shape of the profile, including the smaller morning peak, higher afternoon peak, and lower overnight demand, but small perturbations within an hour period are lost. For instance, the morning energy demand peak occurs at 08:00 - 09:00, but the largest value around 08:30 is underestimated. Similarly, a dip in demand at 11:00 is smoothed out. Not only does the 12-hour downsampled profile not capture small perturbations, but it also misses the key shape of the 5-minute profile. In fact, it only accurately reflects energy demand from 12:00 - 14:00.

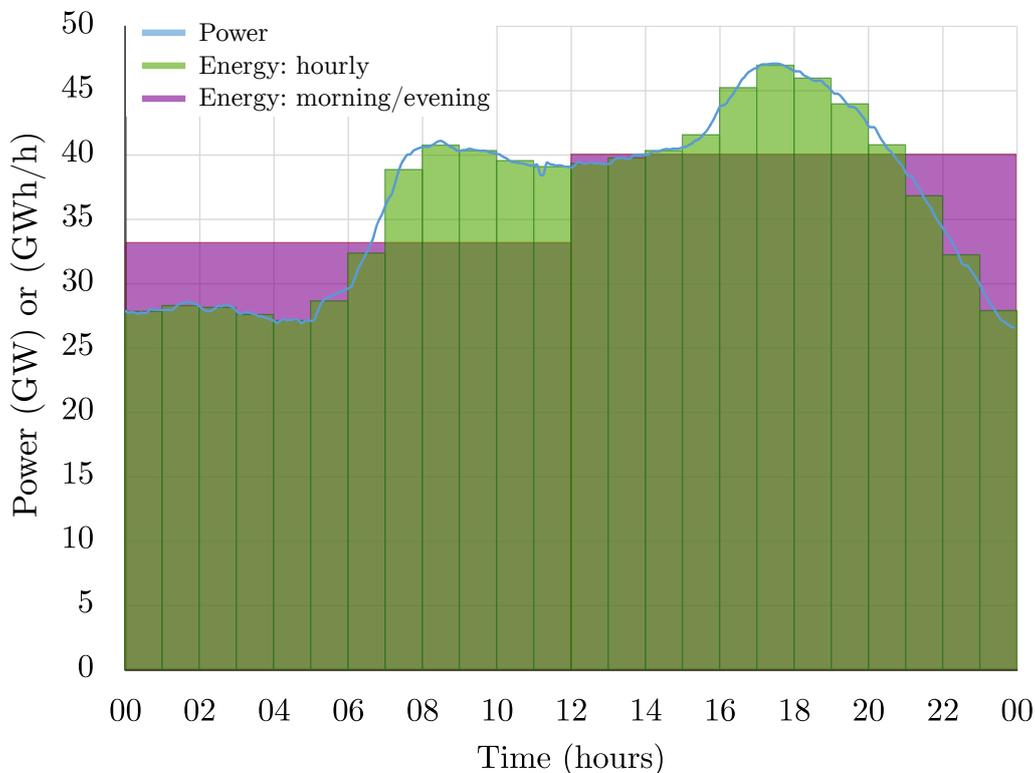


Figure 2.1 UK electricity demand profile for a representative winter day, from <http://www.gridwatch.templar.co.uk/>. Profile is aggregated from 5-minute power consumption (line) to hourly and 12-hourly energy consumption (bars), given per hour for ease of comparison.

When adjusting the temporal resolution in energy system optimisation, Failla *et al.* (2007) showed that downsampling from a 10-minute to a 1-hour resolution led to different results. Yet, several models do downsample, usually from a 1-hour resolution model to a lower resolution. Similar to the aggregation of buildings into nodes, downsampling can be considered as the aggregation of hours into timesteps. Li *et al.* (2016a) considered 2-hour timesteps, while others defined a timestep as four hours (Christidis *et al.*, 2012), six hours (Voll *et al.*, 2015), and twelve hours (Haikarainen *et al.*, 2014; Söderman, 2007; Söderman and Pettersson, 2006). In some studies, the degree of downsampling varies in a day to minimise its effect on the

description of the temporal profile. Weber and Shah (2011) used six timesteps which range in length from one hour to nine hours, with the longest timestep covering the overnight period. The same profiles have since been used in several other studies (Akbari *et al.*, 2016; Mehleri *et al.*, 2012; Zhang *et al.*, 2015a). Over a one year horizon, there are studies which even downsample to monthly timesteps (Majewski *et al.*, 2017; Voll *et al.*, 2013). Considering longer time horizons, Tanaka *et al.* (2017) optimised with one year timesteps over 10 years in their model *DGOPT*, while Cai *et al.* (2009) and Farzaneh *et al.* (2016) both used five year timesteps over 15 and 20 years, respectively.

Most models choose to represent all days in the timeseries by a smaller number of **typical days (TDs)**<sup>1</sup>, instead of, or in tandem with, temporal downsampling. A **TD** will maintain an hourly resolution, but is expected to represent multiple days for which it is typical. When calculating operation costs using 1-hour resolution data, there are still 24 timesteps in a **TD**. Energy generated and consumed in each timestep will be multiplied by the *weight* assigned to the **TD**, equal to the number of days for which it is typical. For instance, Weber and Shah (2011) clustered days based on seasons. Each **TD** used to exemplify a summer, midseason, and winter day represented 92, 153, and 120 days, respectively. By using seasons to cluster, a model will use three or four **TDs** (Koltsaklis *et al.*, 2014; Omu *et al.*, 2013; Wouters *et al.*, 2014; Wouters *et al.*, 2015). Indeed, as well as reducing the time resolution into six timesteps, Weber and Shah (2011) clustered the year into three seasons. Months are also commonly used, leading to twelve **TDs** (Morvaj *et al.*, 2016; Omu *et al.*, 2015).

Seasons and months are subjective clustering choices. More recently, clustering algorithms have been used to select **TDs** from a high resolution timeseries. Kotzur *et al.* (2018a) compared several algorithmic approaches to timeseries aggregation in district energy systems, including ‘hierarchical’, ‘k-means’, and ‘k-medoids’. Instead of selecting days based on subjective similarity, such as months or seasons, days are clustered based on their objective similarities. Each algorithm undertakes a different method of comparing and clustering profiles from each day, although all algorithms aim to minimise error between each day in a cluster. Algorithmic approaches proved more accurate at emulating the full timeseries optimisation results when compared to the average of subjectively chosen clusters. However, the error in objective function value only became satisfactorily small when upwards of 72 **TDs** were used to describe the full, 365 day timeseries.

To ensure ‘peak-load’ days are not lost in clustering, it is possible to retain some days as independent from any typical day. Pfenninger (2017a) referred to these as ‘masked days’ as, once chosen, they are not subjected to clustering. Kotzur *et al.* (2018b) also introduced peak-load days, both in a masked approach and by using the peak-load day in a cluster, instead of the mean or median of all days, as the typical day. In both studies, the introduction of peak-load days improved the accuracy of the reduced timeseries model. Despite not clustering, Voll *et al.* (2013) included two peak-load days to their month timesteps, resulting in a model with 14 timesteps in a year. Mavromatidis *et al.* (2018c), on the other hand, combined both algorithmic clustering and peak-load day masking. This study used ‘k-medoids’ to select 12 **TDs** and two peak-load days representing highest heat and electricity demand.

---

<sup>1</sup>Typical days are also known as representative days in some studies.

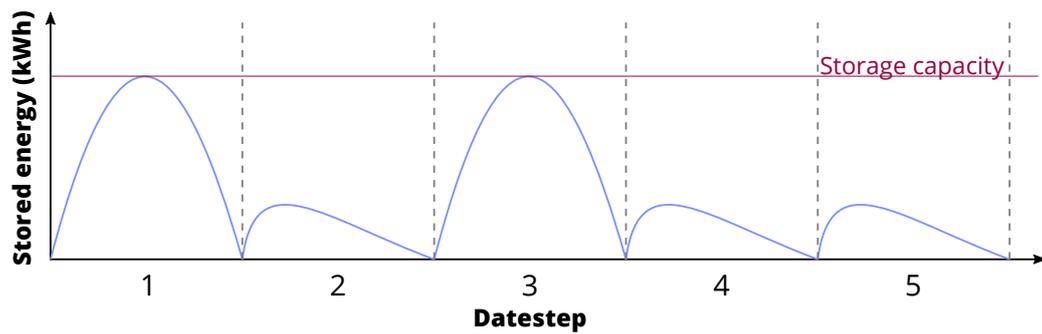
The use of TDs leads to fewer timesteps in a model. When clustering a year into four typical days, Omu *et al.* (2013) reduced the time dimension from 8760 timesteps to 96 timesteps. This order of magnitude dimension reduction will likely have led to greater than an order of magnitude reduction in solution time (Gabrielli *et al.*, 2018; Pfenninger, 2017a). However, the transition between days in the year can no longer be modelled. This does not present a problem for the representation of most technologies, since they are not constrained inter-temporally (i.e. their operation one hour is independent of other hours). However, storage technologies, which are constrained inter-temporally (discussed further on page 20) cannot be operated between TDs. For instance, the stored energy at midnight in TD 2 is not a function of the stored energy at 23:59 in TD 1. This poses a problem for 18 of the studies under consideration in this review, which use both typical days and model storage devices. A common solution is to operate the storage devices on daily cycles, whereby the storage at the beginning of a typical day has to match the storage conditions at the end of that typical day (Gabrielli *et al.*, 2018; Kotzur *et al.*, 2018b; Marquant *et al.*, 2017; Wakui *et al.*, 2014). However, this daily cyclical constraint precludes the modelling of seasonal storage. Where capacity planning is included in the optimisation, no optimal solution would size a storage device larger than required to meet daily storage needs.

To handle seasonal storage, the time dimension could be divided into timesteps and datesteps. Timesteps represent each discrete time interval, retaining the definition assigned to them earlier in this section. Datesteps consider the timeseries at a 24-hour downsampled resolution, equating to 365 datesteps in a year. Timesteps must retain a resolution lower than 24 hours to fit with this datestep definition. Each datestep is assigned to a TD.

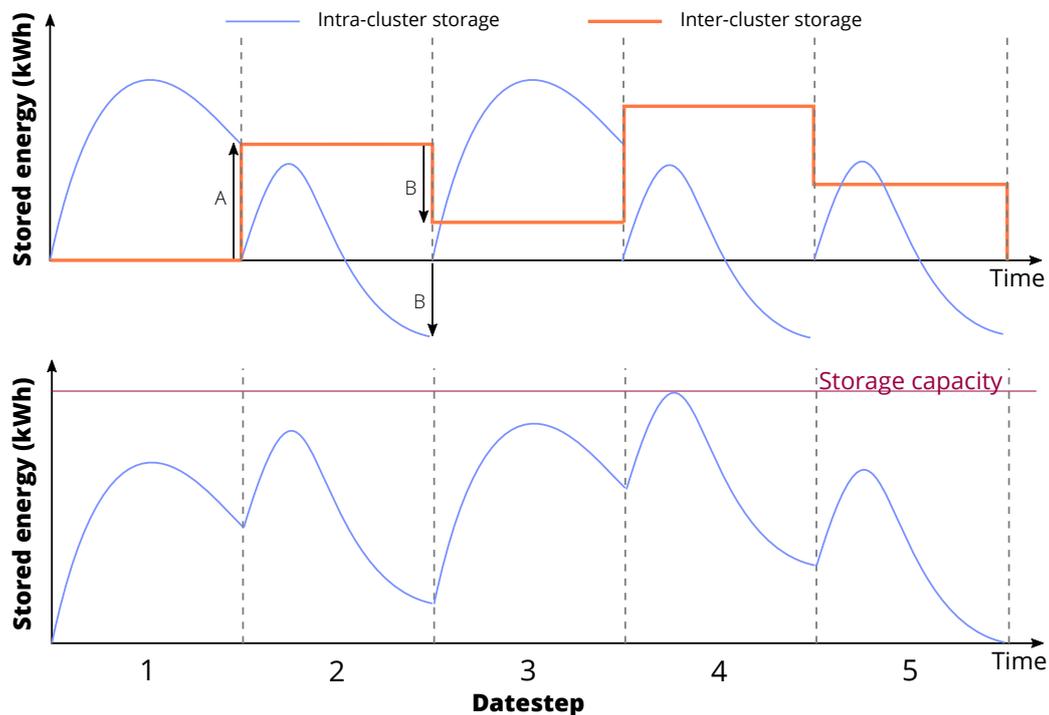
Storage can now be disaggregated into *intra*-cluster storage (for each TD) and *inter*-cluster storage (for each datestep). These two storage types are abstract, only gaining physical significance when recombined by summation. Intra-cluster storage tracks the net flow of stored energy in a particular TD and is a fixed profile for all days associated with that TD. Inter-cluster storage tracks the reservoir of energy across the year and is updated at the end of each datestep based on whether energy has been added (positive intra-cluster storage at the end of the TD), has been removed (negative intra-cluster storage at the end of the TD), or remained constant (zero intra-cluster storage at the end of the TD) across a TD.

Figure 2.2 shows an example five-day storage profile. Days one and three are assigned to TD 1 and days two, four and five are assigned to TD 2. Using only timesteps, the storage must be zero at the end of each day. However, using timesteps and datesteps, intra-cluster storage can be either positive or negative at the end of a typical day, provided the inter-cluster storage does not go below zero. In TD 1, some excess stored energy is available at the end of the day. This spills over into the next datestep, allowing TD 2 to draw more energy from storage than would otherwise be possible. The maximum capacity of the storage vessel is given as the maximum sum of inter- and intra-cluster storage associated with a TD. The combination of these two timeseries dimensions has been tested in two methodological studies (Gabrielli *et al.*, 2018; Kotzur *et al.*, 2018b). In both studies, the error in storage capacity reduced when using TDs compared to using the full timeseries. There is a time penalty incurred relative

to the use of daily cyclic storage, although the penalty is still at least an order of magnitude lower than modelling with the full timeseries.



(a) Daily cycle storage.



(b) Inter-cluster storage.

Figure 2.2 Example of a) daily cycle storage and b) inter-cluster storage when using TDs to describe a full timeseries. Daily cycle storage requires that the storage vessel is fully discharged at the end of each day and that all days within the same cluster operate identically. Inter-cluster storage separates the storage vessel into two abstract vessels: intra-cluster storage and inter-cluster storage. Intra-cluster storage can be negative or positive (or zero) at the end of each day, but must have the same profile on all days within the same cluster. The inter-cluster storage vessel only changes at the start of each day in the timeseries. It is independent of clustering and is respectively charged or discharged by the end-of-day state of the intra-cluster storage. The upper trace in b) shows the two abstract vessels, the lower trace shows their superposition, which is the physical operation of the storage vessel.

When optimising only for technology operation, an hourly time resolution is standard. There are no cases of time resolution lower than this and Good and Mancarella (2017) even increased the resolution to 30 minute timesteps. However, the time horizon of operation models is shorter than in planning models. Only Capuder and Mancarella (2014) and Orehounig *et al.* (2015) run the operation optimisation over an entire year. A one day time

horizon is far more common (Bischi *et al.*, 2014; Bucciarelli *et al.*, 2018; Cesena and Mancarella, 2018; Good and Mancarella, 2017; Pazouki *et al.*, 2014; Reddy, 2017; Vahid-Pakdel *et al.*, 2017; Yokoyama *et al.*, 2014; Zheng *et al.*, 2018). To consider a longer time horizon whilst maintaining a small time dimension, Wakui *et al.* (2014) introduced TDs to operation schedule optimisation, selecting one for each month in the model.

There is always a fixed limit to the length of the time horizon; a system cannot be modelled ad infinitum. Usually, planning models consider a one year horizon, ignoring inter-annual variations in parameters. Although temporal resolution is heavily reduced, models which include multi-year time horizons show clear inter-annual variation in parameters and results (Cai *et al.*, 2009; Farzaneh *et al.*, 2016; Tanaka *et al.*, 2017). End-user demand increases from year to year, almost doubling in Delhi over 20 years (Farzaneh *et al.*, 2016) and requiring a three-fold increase in technology capacity over 20 years in Tokyo (Tanaka *et al.*, 2017). Models that use a one year time horizon in planning technology capacity will often consider that year as a ‘typical’ year, similar to TDs (Mavromatidis *et al.*, 2018b; Wakui *et al.*, 2014; Yokoyama *et al.*, 2016). The resulting optimal technology portfolio is thus valid for the given typical year, yet ignores how the supply infrastructure will change over a lifetime of 15 (Capuder and Mancarella, 2014), 20 (Wouters *et al.*, 2015), or even 40 (Marquant *et al.*, 2017) years. Additionally, in considering a year as ‘typical’, storage capacity must be modelled accordingly. The storage at the beginning of the year must equate to the storage at the end of the year. Without cycling the storage, if there is additional stored energy of  $Y$  units at the end of a year, that would balloon to  $20Y$  units over a 20 year period. However, ‘typical’ year storage capacity will not have factored in the additional capacity required to ensure a  $20Y$  surplus after 20 years.

## 2.1.2 Energy technologies

The supply portfolio of a model comprises energy technologies, which may be exogenous or endogenous within a model. Exogenous supply portfolios are used in optimisation when designing an operation schedule for a known system (e.g. Good and Mancarella, 2017; Vahid-Pakdel *et al.*, 2017), or when the portfolio is limited to a small number of scenarios (Capuder and Mancarella, 2014; Orehounig *et al.*, 2015; Wouters *et al.*, 2014). In most cases, the supply portfolio is endogenous, and plays a major role in the purpose of a given study, namely to make purchase options for a district to meet demand whilst minimising a given objective. The number of possible technologies can vary, from just two (Keirstead *et al.*, 2010; Reddy, 2017; Söderman, 2007) to over ten (Farzaneh *et al.*, 2016; Kotzur *et al.*, 2018b), but most consider around five to eight technologies. Factors such as which technologies to purchase and their capacity, location, and operational dispatch must all be considered in many models.

### Conventional technologies

Conventional energy technologies are those which are most frequently found in existing energy systems. They meet the demand of only one energy carrier and are generally considered to be completely dispatchable, i.e. available to meet energy demand at all times. Thermal demand and electricity demand are conventionally met by geographically opposing

technologies. Conventional thermal technologies are sited within a building, while electricity demand is met by a national electricity grid, where energy generation is geographically far from points of consumption.

A **natural gas boiler (NGB)** is the conventional heat provision technology. **NGBs** are represented in models as energy supply technologies which consume a fuel at a given efficiency to provide heat (Figure 2.3a). Although natural gas is the commonly assigned fuel, biomass, oil, wood, and electric boilers are also studied (Kotzur *et al.*, 2018a; Kotzur *et al.*, 2018b; Mavromatidis *et al.*, 2018a; Mavromatidis *et al.*, 2018c; Omu *et al.*, 2013; Omu *et al.*, 2015; Wakui *et al.*, 2014; Yazdanie *et al.*, 2016). The efficiency by which they operate depends on the study, ranging from 70% (Jennings *et al.*, 2014) to 95% (Buoro *et al.*, 2014).

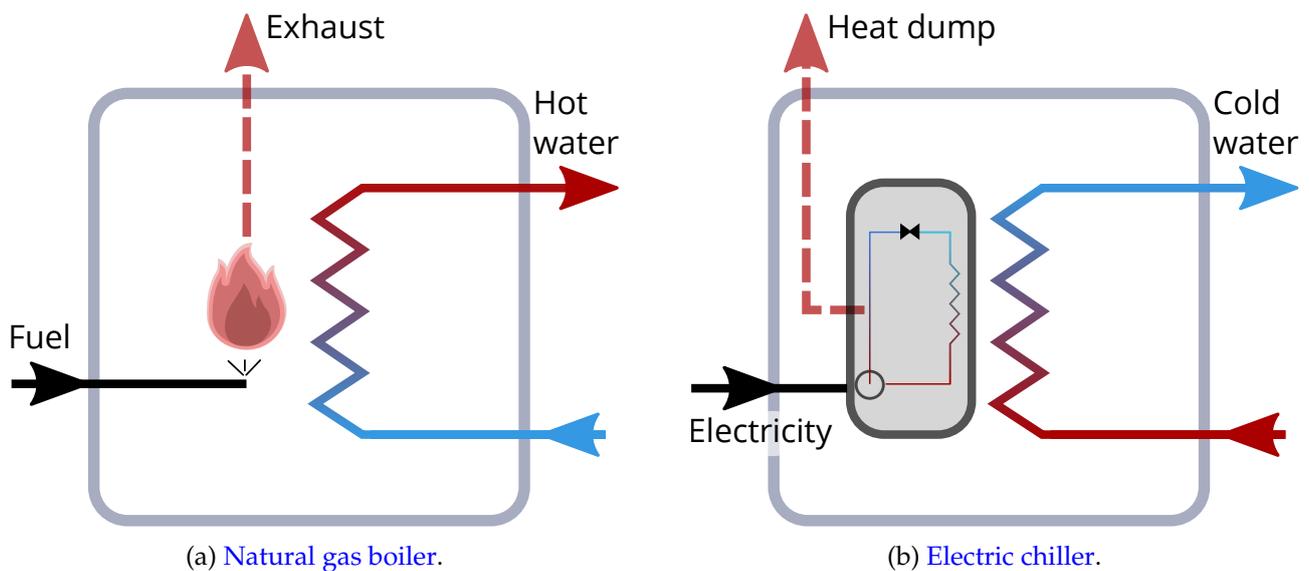


Figure 2.3 Graphical representation of conventional energy technologies, as depicted in linear energy system models. Heat dump and exhaust energy streams refer to lost energy, quantified in models using technology efficiency.

Cooling demand is typically met by an **electric chiller (ECh)**, sometimes referred to as an electric heat pump. They are represented similarly to boilers, but a secondary loop is required to describe the medium to which heat is dumped when producing cold air or water (Figure 2.3b). Usually, this medium is air and is described by an **air source heat pump (AHP)** operating in cooling mode. Cooling **AHPs** employ rooftop units to dump heat directly into the local atmosphere (Zmeureanu, 2002). Water and the ground are also possible heat dumping media. **Ground source heat pumps (GSHPs)** capitalise on the relatively static, low ground temperature and are also frequently included in models (Jennings *et al.*, 2014; Mavromatidis *et al.*, 2018a; Omu *et al.*, 2015). Instead of considering technology efficiency, which requires an understanding of heat transfer to the dump medium, cooling technologies use the **coefficient of performance (COP)** as measure of technology performance. The **COP** is defined as the technology thermal output relative to electrical energy input. **COP** values above 1 are expected, so should not be confused with technology efficiency, for which an efficiency greater than 1 (100%) would be thermodynamically impossible. As with boiler

efficiency, there is little agreement on chiller COP. Values range from 2.5 (Buoro *et al.*, 2014) to 5.02 (Yokoyama *et al.*, 2014).

As districts are most prominent in cities, access to [national grid electricity \(GridE\)](#) appears in most studies. This electricity can be purchased from the grid, and it is possible to sell excess electricity back to the grid; it is concurrently an infinite source and infinite sink of electricity within the model. For this, a technology is rarely modelled. Nevertheless, intricacies of AC power flow are included by some studies. Where electricity is the only energy carrier, studies often revert to an electrical power system representation. This includes a substation with a power factor and power transformation capabilities (Bucciarelli *et al.*, 2018; Reddy, 2017; Tanaka *et al.*, 2017). In linear multi-energy district models, substations are not modelled. In some studies, local generators are modelled instead of the grid. Diesel generators are the incumbent method for meeting demand in an electrical system representing Grenada (Garcia and Weisser, 2006) and in an illustrative U.S. district (Hoke *et al.*, 2013).

### Multi-energy technologies

Most studies create models to investigate the possibility of lower dependence on conventional technologies. One method involves the use of waste heat to meet both heat and cooling demands. The waste heat usually comes from the exhaust of an electricity generator, or prime mover, creating a [CHP](#) technology from conventional technologies such as [Stirling engines \(SEs\)](#), [gas turbines \(GTs\)](#) or [internal combustion engines \(ICEs\)](#), or from the novel use of [fuel cell \(FC\)](#) technologies (Figure 2.4a). The available heat from a [CHP](#) technology depends on its electrical efficiency and is related to the electrical power output using the [heat-to-power ratio \(HTP\)](#). If a technology has a higher electrical efficiency, its heat output and [HTP](#) will inevitably be lower. Additionally, not all the waste heat from a [CHP](#) technology will be useful, due to its low temperature, or will be lost in heat transfer, leading to [CHP](#) technologies which have ‘overall’ efficiencies of 80% to 90% (Zhang *et al.*, 2015a).

Depending on the difference between heat and electricity demand in a district, certain technologies will prove more suitable than others. Zhang *et al.* (2015a) found that a relatively high heat demand favours the use of [ICE](#) or [SE](#) technologies, which both have a relatively low electrical efficiency and reciprocally high [HTP](#). Similarly, Koltsaklis *et al.* (2014) found that there was a greater probability of a [SE](#) or [ICE](#) being selected for installation in a district when optimising 1,000 Monte Carlo generated scenarios. However, in most scenarios studied by Zhang *et al.* (2015a) there was a sufficiently low heat demand to warrant investment in a proton exchange membrane [FC CHP](#), due to its relatively high electrical efficiency of 40%. Distinguishing different [CHP](#) technologies is not common practice, however, with only a few studies choosing to do so (Bischi *et al.*, 2014; Mehleri *et al.*, 2012; Wouters *et al.*, 2014). Therefore, the prime mover is not specified, leaving only the fuel input, technology efficiency, and [HTP](#) for a generic [CHP](#) to be defined.

For exhaust heat to produce cooling, a [heat recovery absorption refrigerator \(HRAR\)](#) is required. Similar to an electric chiller, a [HRAR](#) depends on a refrigerant cycle to provide cooling. In an electric chiller, a gaseous refrigerant is compressed. This high pressure refrigerant vapour is then condensed by transferring energy to an external source: the heat

dump medium. The pressure of the resulting liquid refrigerant is reduced, by use of an expansion valve, before re-vapourising by drawing heat from the space cooling water flow. However, in a [HRAR](#), a high pressure refrigerant vapour can be prepared by applying heat to a water-refrigerant solution (Srikhirin *et al.*, 2001). In doing so, electricity is replaced by heat as the primary energy source for the technology (Figure 2.4b). Although electricity is required to drive a pump in the [HRAR](#), the energy consumption is generally considered negligible (Srikhirin *et al.*, 2001).

While a [HRAR](#) is not a multi-energy technology, it is often considered as part of a single [CCHP](#) (Figure 2.4c) which produces electricity, heat, and cooling simultaneously. In the same vein, a [CCHP](#) can be formulated by the inclusion of an [ECh](#) powered by the electrical output of a [CHP](#). Ameri and Besharati (2016) included both an [ECh](#) and a [HRAR](#) in their model of a [CCHP](#). They found that the inclusion of both technologies ensures greater overall use of the multi-energy system, with the [HRAR](#) providing baseload cooling, topped up by the [ECh](#) at peak loads. However, the [ECh](#) often operated conventionally, drawing directly from the grid. In contrast, Majewski *et al.* (2017) placed a [HRAR](#) at each node, powered by a [CHP](#)-fed district heating system. The whole system acts as a [CCHP](#), but cooling is decentralised and heat and electricity are centralised. Omu *et al.* (2015) found that it is more cost-optimal to invest in a [HRAR](#) than an [ECh](#), if a model must rely on a [CHP](#). But, if a model is free to choose whether to invest in a centralised system, Wouters *et al.* (2015) found that [EChs](#) were more favourable than [CHP](#) driven [HRARs](#) for an illustrative Australian district. Most studies which include cooling demand allow the installation of an [ECh](#) or a [HRAR](#). Uniquely, Ren *et al.* (2010) only modelled a [HRAR](#) as being able to meet cooling demand, yet direct gas firing was another possible heat source, which allowed the [HRAR](#) to be disconnected from the [CHP](#).

As with boilers, the most common fuel source for a modelled [CHP](#) is natural gas. Yet, there are exceptions. Orehounig *et al.* (2015) and Zheng *et al.* (2018) model biomass fuelled [CHPs](#), in which the [CHP](#) contributes heavily to meeting system demand. Henning *et al.* (2006) even modelled a [CHP](#) with seven fuel sources: biomass, waste, coal, rubber, plastics, oil, and animal fat, concluding that waste as the primary fuel source was cost optimal. Waste heat does not always need to be a by-product of electricity generation. Koltsaklis *et al.* (2014) incorporated waste heat from a refinery into their model, allowing it to be purchased in favour of [CHP](#) generated heat. The optimal solution capitalises on this cheap waste heat source, particularly for meeting demand in buildings with a relatively high heat demand.

## Renewable energy

Another method to reduce dependence on conventional technologies is to replace them with local-scale renewable technologies. Three renewable energy technologies are included in district-level models: [wind turbines \(WTs\)](#), [solar photovoltaic panels \(PVs\)](#) and, to a lesser extent, [solar thermal panels \(STs\)](#). These technologies are unique in the way they consume primary energy to produce electricity or heat. Instead of assuming an infinite source of energy from a fuel, the primary energy varies both temporally and spatially. The small spatial scale of a district system means that the latter variation has little to no effect and is not modelled

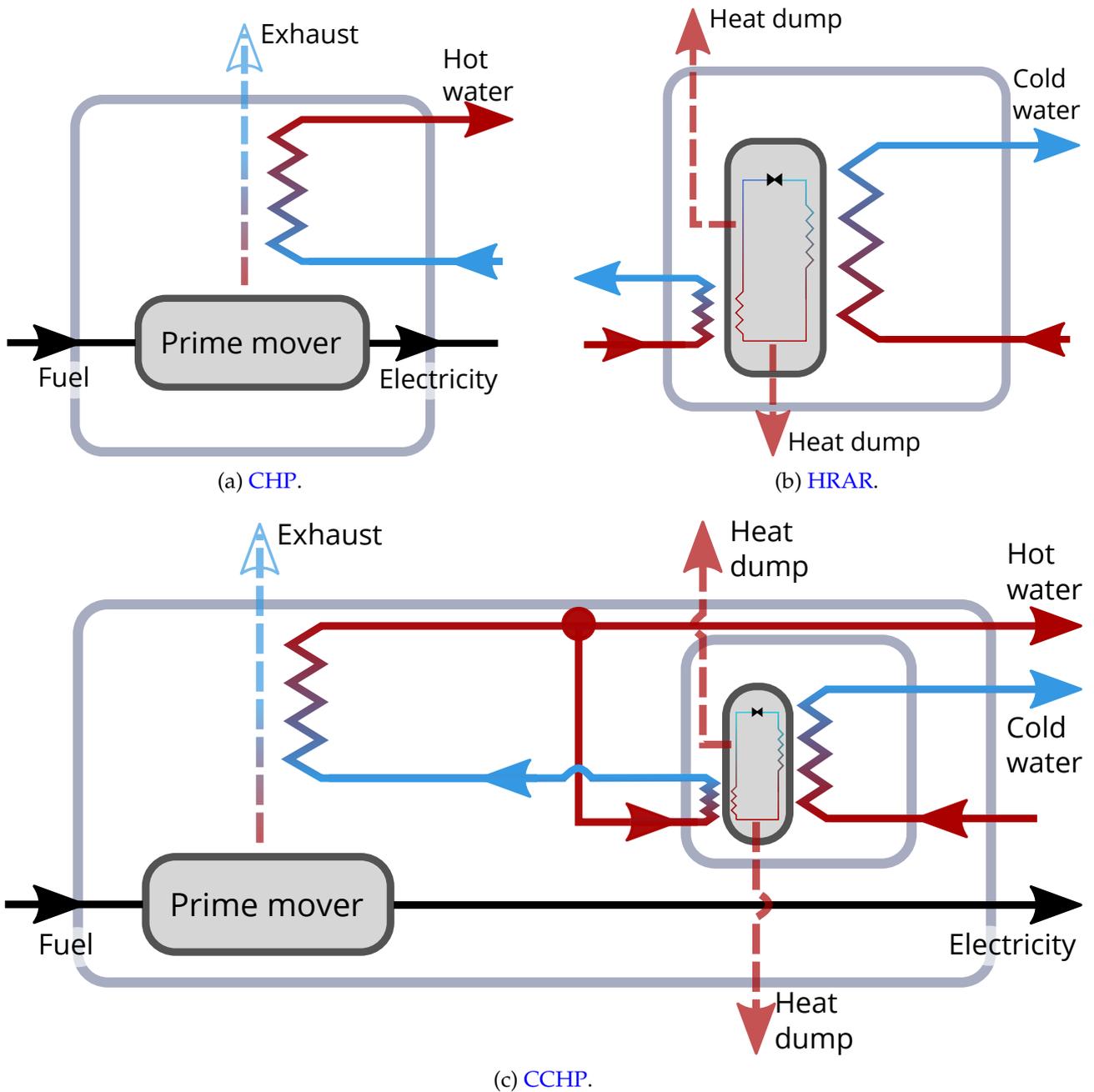


Figure 2.4 Graphical representation of multi-energy technologies, as depicted in linear energy system models. Heat dump and exhaust energy streams refer to lost energy, described by technology efficiency. Prime movers in (a) and (c) refer to any device for producing work from fuel, such as a [SE](#).

in district-level studies. However, handling timeseries variation of renewable energy is the primary aim of many of the reviewed studies.

In just one day, the availability of a renewable energy resource can vary significantly. This availability may not match periods of higher demand, requiring additional technologies to reconcile the temporal mismatch. Garcia and Weisser (2006) modelled the possibility of WT penetration into Grenada, by coupling the variable energy source with hydrogen storage (StoreH2) and a hydrogen FC. Indeed, the optimal coupling of intermittent renewable energy with storage technologies is studied frequently (Bucciarelli *et al.*, 2018; Carpaneto *et al.*, 2015; Hoke *et al.*, 2013; Tanaka *et al.*, 2017).

Renewable energy availability varies throughout the day, but when modelling it is important to understand that its value at any given time is stochastic. Wouters *et al.* (2015) optimised an Australian district for a fixed solar irradiation, but showed through Monte Carlo simulation that the actual output of the installed PVs would vary sufficiently to impact the optimal result. Although the optimal technology capacity would not noticeably change, its reliance on consumption of on-site production would significantly increase with a reciprocal decrease in the export of electricity to the grid. Unlike the ex-post analysis of this stochasticity analysed by Wouters *et al.* (2015), many studies have incorporated renewable energy uncertainty into the optimisation. Pazouki and Haghifam (2016) modelled wind energy with a  $\pm 10\%$  variance in a scenario optimisation model. Other approaches have included the application of a Weibull distribution (Reddy, 2017; Vahid-Pakdel *et al.*, 2017), a Gaussian process (Zhou *et al.*, 2018), and direct sampling of historical data (Tanaka *et al.*, 2017) prior to risk-aware optimisation.

## Energy storage

When incorporating renewable energy and multi-energy technologies into district systems, energy storage becomes a useful and sometimes necessary component of a model. Increasing use of intermittent renewable resources within the supply portfolio leads to less control of when energy is produced, thereby requiring storage as a buffer. Equally, since CHPs produce heat and electricity simultaneously, storage can decouple energy streams, if heat and electricity demand do not coincide. Storage can also allow technologies to operate at a continuous load instead of cycling to meet varying demand. As a result, supply technologies can feasibly capitalise on higher efficiency part-load ranges, extend technology lifetime, and reduce maintenance requirements (Martin and Thornley, 2013; Powell *et al.*, 2013).

In multi-energy systems, balancing energy supply and demand is often achieved by use of thermal energy storage (StoreT). Morvaj *et al.* (2016) saw an increase in StoreT when modelling scenarios involving greater PV investment. Given the non-dispatchable nature of the PV, a CHP was forced to meet electricity demand at times of low PV output. This led to waste heat production at times of low heat demand and therefore required StoreT. Li *et al.* (2016a) used the same StoreT to store energy for heating and cooling, in a system with a CCHP and PVs. In all seasons, thermal demand peaks were reduced by use of StoreT, although heat and cooling demand never coincided. Similarly, Wang *et al.* (2015) found that StoreT was a useful asset, particularly when system heat demand was met with an increased

contribution from either a **CHP** or a **ST**. Furthermore, when **electrical battery storage (StoreE)** and **StoreT** are included in a model, they can both be utilised to balance the output of a **CHP** that must run during the troughs of non-dispatchable renewable output (Hawkes and Leach, 2009). As well as balancing energy supply and demand, Christidis *et al.* (2012) showed that it is possible to capitalise on electricity price fluctuations, if **StoreT** is coupled to a **CHP**. The **CHP** can operate when the export of electricity is most lucrative, storing waste heat for meeting space heat demand at other times of the day.

Thermal energy is usually stored as a liquid in an insulated vessel. As such, the available energy in the vessel is a function of its size and the liquid temperature. By considering the volume of a storage vessel, Wang *et al.* (2015) saw that the actual energy storage capacity of the vessel changed depending on the time of year. In summer, a lower temperature difference between the stored heat and the surrounding environment led to lower effective stored energy than in winter. Storage vessels are, however, usually described by their energy capacity (e.g. Carpaneto *et al.*, 2015; Li *et al.*, 2016a; Morvaj *et al.*, 2016) or are assumed to operate consistently at the same temperature, effectively statically linking volume and energy (Buoro *et al.*, 2014). To introduce thermodynamic effects, storage vessels are given a percentage loss related to their ‘self discharge’ rate. Gabrielli *et al.* (2018) applied a 0.5%/h storage loss on top of a  $0.001f(T)$  loss, depending on the external temperature  $T$ . The influence of external temperature was calculated ex-ante and applied exogenously by both Gabrielli *et al.* (2018) and Wang *et al.* (2015). The modelled self discharge of **StoreT** does vary depending on the study, ranging from rates of 0.1%/h to 5%/h (Carpaneto *et al.*, 2015; Gabrielli *et al.*, 2018; Kotzur *et al.*, 2018b; Mavromatidis *et al.*, 2018b; Vahid-Pakdel *et al.*, 2017). **StoreEs** also suffer from self-discharge rates, but they are set lower, at around 0.05% (Kotzur *et al.*, 2018a) to 0.1% (Mavromatidis *et al.*, 2018b).

Storage vessel charge and discharge are modelled similarly to energy consumption and production of supply technologies, including an efficiency to account for losses as high as 10% for providing energy to the system from either **StoreE** or **StoreT** (Mavromatidis *et al.*, 2018b; Mavromatidis *et al.*, 2018c; Pazouki and Haghifam, 2016; Weber and Shah, 2011; Zhou *et al.*, 2018). Additionally, charging and discharging are often limited to a percentage of the full capacity of the storage device. By applying a maximum 50% discharge limit, Wang *et al.* (2015) dictated that the modelled **StoreT** must discharge over at least two hours. Discharge limits are placed on **StoreTs** in other studies, ranging from 40% (Mavromatidis *et al.*, 2018b) to 60% (Li *et al.*, 2016a). Wouters *et al.* (2015) discounted **StoreT** limits, choosing to only limit the charge and discharge of **StoreE** to 90% and 85% of its capacity, respectively.

Another storage type, **StoreH2**, is relatively bespoke, but is modelled in both electricity-only districts and multi-energy districts. Al Rafea *et al.* (2017) modelled the use of excess power from intermittent renewable technologies to generate hydrogen by electrolysis for storage. When sufficient hydrogen has been accumulated, they suggest that it can be ‘injected’ into the national natural gas grid, in which the hydrogen is blended back into the national gas network. Indeed, blending hydrogen with natural gas is becoming a more prominent method for handling hydrogen production by electrolysis, with most modern appliances being capable of burning blends up to 20% hydrogen, by volume (Quarton and Samsatli, 2018).

Disconnected from national infrastructure, Garcia and Weisser (2006) modelled [StoreH2](#) as a way of handling intermittent wind power generation, but the cost of the technology was prohibitive at 2006 prices. However, on optimising for a 2010 scenario, the expected cost reduction was sufficiently favourable to make a [StoreH2](#) solution viable, while also decreasing system CO<sub>2</sub> emissions from the 2006 optimisation by 10%. [StoreH2](#) is often modelled as a seasonal storage device in multi-energy systems (Gabrielli *et al.*, 2018; Kotzur *et al.*, 2018b). When seasonal variations are considered, [StoreH2](#) is more heavily relied upon than either [StoreE](#) or [StoreT](#).

As discussed on page 12, storage vessels are usually only modelled to operate on a daily cycle, due to the use of [TDs](#) following time dimension aggregation. This is apparent in the operation schedules given by Li *et al.* (2016a) for three [TDs](#), while other studies recognise that seasonal storage is not feasible following their temporal adjustments (Marquant *et al.*, 2017; Mavromatidis *et al.*, 2018b; Wouters *et al.*, 2014). To model seasonal storage, in which energy is stored in one season for use in another, the model must be described by the full time series or the inter-cluster storage technique described in Figure 2.2 must be used. Gabrielli *et al.* (2018) used the latter approach, finding that although [StoreE](#) remained in a daily or weekly cycle, [StoreT](#) and [StoreH2](#) capitalised on seasonal variability in demand to charge over summer and autumn, then discharge over winter and spring. Kotzur *et al.* (2018b) showed that only 12 [TDs](#) were required to converge on the correct [StoreH2](#) capacity, providing those [TDs](#) were algorithmically chosen and inter-cluster storage was implemented. On conducting a full year optimisation Buoro *et al.* (2014) also found that stored energy would follow a seasonal cycle, requiring [StoreT](#) eight times larger when compared to optimising for just one week.

## Distribution

District energy systems with several demand nodes require distribution networks. These may be electrical lines, thermal pipes, or natural gas pipes. Similar to supply and storage technologies, distribution may be endogenous or exogenous in a model. If endogenous, the available routes are defined in most models (e.g. Haikarainen *et al.*, 2014; Omu *et al.*, 2013; Weber and Shah, 2011) and, as discussed previously, this routing of distribution networks tends to follow roads. Where routes are undefined, the district is given to be a new development with no existing road infrastructure to follow (Keirstead *et al.*, 2010; Morvaj *et al.*, 2016). One exception is Ameri and Besharati (2016), who did not consider the road layout when connecting buildings in an existing district. Instead, a straight line connection between demand nodes was assumed.

When energy flows between two geographically separated nodes, losses to the local environment are noticeable. Districts may be small enough to neglect electrical line losses (Hoke *et al.*, 2013), but thermal energy losses are likely to be incurred. However, it is not common for pressure and temperature to be tracked in a linear district network. Khir and Haouari (2015) uniquely modelled cold water flow in the network as a secondary [mixed integer linear programming \(MILP\)](#) model overlaid onto the technology capacity planning model. Changes in pipe pressure and temperature were linearised along the length of the network. The initial pressure level was set by a compressor at the chiller plant, with the

pressure needing to be high enough leaving the compressor to ensure adequate pressure at the end of the network. Similarly, the chiller plant was able to undercool the water, such that it maintained a sufficiently low temperature at the end of the network. Although Khir and Haouari (2015) provided no indication of the change in result, the computational time required to undertake such complex modelling increased by up to three orders of magnitude. Haikarainen *et al.* (2014) compromised on the complexity of tracking pressure, by ensuring that water flow in any section of pipe could overcome pressure losses associated with a particular flow velocity and pipe diameter. Compressors were made available at every node, and their required energy input was calculated using a linear relationship between flow velocity and pipe diameter. No attempt was made to maintain or track temperature, however. Instead, temperature losses were represented by a percentage energy loss per unit pipe length.

To avoid increased complexity, it is common to ignore both system pressure and temperature. Most studies only apply a percentage loss to the energy being distributed in the network (Mavromatidis *et al.*, 2018b; Morvaj *et al.*, 2016; Orehounig *et al.*, 2015; Weber and Shah, 2011) or apply no loss at all (Li *et al.*, 2016a; Omu *et al.*, 2013). Where energy losses are considered, they are usually given per unit pipe length and constitute 5% - 10% of the heat distributed in the network (Haikarainen *et al.*, 2014; Weber and Shah, 2011), although Taljan *et al.* (2012) found that an existing district in Australia had recorded average network losses of 24%.

At each node in a district, distribution technologies can either supply or consume an energy carrier, depending on the direction in which they are transporting energy. As such, they are usually considered as two linked technologies, acting similarly to a supply or demand technology. Energy that is consumed in one node is then instantly available to be supplied at another, connected node. Weber and Shah (2011) also included an energy cascade equation in their representation of the network, to ensure that heat provided to the network is at or above a required temperature. At each cascade level, temperature is fixed, which allows mass flow to be tracked linearly relative to energy flow. Christidis *et al.* (2012) calculated mass flow rates ex-ante for the Berlin-wide heating network, based on known, temporally varying temperatures. Pump power requirements throughout the network could then be calculated, using the known flow rates.

Akbari *et al.* (2014) noted that if any distribution lines are bi-directional, or form loops when connected to specific nodes, an unrealistic dissipation of heat may occur. Heat could be extracted from a node then recirculated back to the same node in the same timestep, but with lower energy content due to heat dissipation from the pipe. It may prove optimal to dissipate this heat if, for instance, a CHP would benefit from producing more electricity without a suitable source of heat demand. Instead, several studies link nodes in the district heating network such that no node can be revisited, i.e. flow must be unidirectional along all pipelines<sup>2</sup> (Akbari *et al.*, 2014; Li *et al.*, 2016a; Marquant *et al.*, 2017; Mehleri *et al.*, 2012; Morvaj *et al.*, 2016; Wouters *et al.*, 2015). The distributed energy must flow only in one direction, based on

<sup>2</sup>This approach has been likened to the 'travelling salesman' problem (Akbari *et al.*, 2014), a classic optimisation problem in which the shortest path to connect a number of points is sought, finishing at the starting point without revisiting any intermediate point.

the order in which the nodes are linked. This approach ignores the return flow of heat, but as the return flow holds no 'useful' energy it is reasonable to discount it from the model.

Few studies which model the distribution network as endogenous conclude with the existence of a full-scale district network being optimal. Ameri and Besharati (2016) found that only interconnecting five of the seven nodes with thermal pipelines is considered optimal by their model. By decreasing the CO<sub>2</sub> emissions allowed in the system, Morvaj *et al.* (2016) showed that greater interconnection of distribution lines was possible. Conversely, with no emissions constraints, one or two buildings were found to interconnect, but in their own separated networks. All buildings in a UK 39-node network were connected together in the optimal result given by Weber and Shah (2011), in which no emissions constraints were in place. However, it is likely that distribution costs were not considered. The 750 - 1000 GBP/m cost of heat distribution pipework used by Omu *et al.* (2013) led to no district heating network being installed in a cost-optimal solution.

### 2.1.3 Energy demand

In district-level optimisation, energy demand is exogenous and assigned to a building or node. The difficulty in finding an optimal technology portfolio is often directly caused by the need to meet different energy demands, which vary both temporally and spatially. This may be exacerbated by the introduction of artificial complexity in the nodes which occupy the district. For instance, Hawkes and Leach (2009) considered a three building district including a hospital, hotel, and leisure centre, while Zhang *et al.* (2015a) modelled a school, a hotel, a restaurant, an office building, and a residential building. Each building in the district has a unique energy demand profile and peak demand, and may also include simultaneous cooling and heat demand (e.g. hospital, hotel, office). It is unlikely that these districts exist, but designing a supply and distribution system for them is sufficiently complex to warrant computational optimisation. Nevertheless, where buildings are heterogenous, the demand may still be relatively homogenous. Omu *et al.* (2013) modelled a district including a café, office, residential space and workshops. Their electricity demand profiles had very similar shapes, albeit with different peak loads. The shape of heat demand profiles were more variable, however.

Some districts have more homogenous end-use. Residential areas are a particularly common example of this, where nodes in the network have exactly the same profile shape, albeit with a different peak demand (Ameri and Besharati, 2016; Wouters *et al.*, 2015; Zheng *et al.*, 2018). Orehounig *et al.* (2015) also modelled a residential area, but took into account the age of different buildings, thereby producing different heat demand profiles. However, since occupants follow the same energy use patterns, these buildings had exactly the same electricity profile shape. Buoro *et al.* (2014) modelled an industrial site, but noted that demand between buildings will differ, depending on the production taking place in each building. Although Koltsaklis *et al.* (2014) modelled the interaction between an industrial and residential area, only the residential area had heat demand since the industrial area was a source of waste heat.

A good knowledge of the spatio-temporal variations in building energy demand is required when preparing an optimisation model. If the district consists of existing buildings, historical demand data may be used directly (Buoro *et al.*, 2014; Jennings *et al.*, 2014). Where no historical data exists, such as for new buildings, demand can be simulated (Li *et al.*, 2016a; Mavromatidis *et al.*, 2018a; Morvaj *et al.*, 2016). Indeed, the SynCity optimisation model developed by Keirstead *et al.* (2010) could simulate demand. Although this would suggest an optimisation model with endogenous demand, the internal workflow actually led to fixed demand prior to optimisation. However, simulations require detailed knowledge of the buildings. When a district comprises several types of heterogeneous buildings, the quantity of required data and man-hours may be prohibitive (Reinhart and Cerezo Davila, 2016). Additionally, they are difficult to validate. For instance, simulated and actual demand of commercial buildings in the UK may exhibit discrepancies of 16%–500% (The Carbon Trust, 2011). A more favoured approach in district energy optimisation is to use archetypal or reference demand profiles from relevant buildings as representative of future energy demand for the district (Fonseca and Schlueter, 2015; Mehleri *et al.*, 2012; Omu *et al.*, 2013; Swan and Ugursal, 2009).

## 2.2 Shortcomings of existing models

The previous section presented the existing paradigm in modelling district energy systems. However, there are several pitfalls to these methods. Namely, the models aggressively simplify the problems, place too much weight on the validity of deterministic results, and are not sufficiently transparent such that they can be easily reproduced. Each of these shortcomings will be discussed in this section, alongside examples of state-of-the-art methods which have recently been studied to address them.

### 2.2.1 Model simplifications

To derive useful information from energy system optimisation, the limitations of the approach must be understood. Linear models are often chosen to represent districts because they can solve a problem in a reasonable amount of time. They also deterministically achieve a ‘global’ optimum, giving a modeller the best possible solution for the provided inputs. Inevitably, the quality of data inputs has a profound effect on the optimum solution. To receive a realistic result from optimisation, the inputs themselves must be as good a representation of reality as possible. However, reality is often nonlinear and thus cannot be represented in a linear model, even if the data is available.

For instance, due to the complex processes used by most technologies to convert and store energy, their efficiency profiles tend to be nonlinear and dependent on several externalities. Performance curves of technologies such as CHPs, chillers, and heat pumps can be quadratic (Basu, 2013; Wu *et al.*, 2012) or even cubic (Best *et al.*, 2015; Ikeda and Ooka, 2016). Metaheuristic algorithms can handle these complexities, but linear algorithms cannot. This limited the scope of the RESCOM MILP model, in which commercial properties were not

modelled due to worries that cooling systems are too nonlinear (Jennings *et al.*, 2014). More often, however, existing linear models ignore this complexity by using only the nominal or reference efficiency of a technology, i.e. the efficiency at full load (e.g. Buoro *et al.*, 2014; Carpaneto *et al.*, 2015; Koltsaklis *et al.*, 2014; Zhang *et al.*, 2015a).

The process of storing thermal energy does not require energy conversion, but has its own complexities. When water is stored in a tank, it is usually stratified (Nielsen, 2005). Hot water enters the top of the tank and cold water is extracted from the bottom of the tank. There exists a region in these tanks in which a temperature gradient exists between the hot and cold regions, considered to be unused from a useful energy perspective (Campos Celador *et al.*, 2011; De Césaró Oliveski *et al.*, 2003). To represent this region, the tank is usually discretised into vertical layers, between which differential equations relating mass and energy flows can be resolved. It is agreed that a 1-D representation of the tank is possible, as temperature gradients in the horizontal plane are insignificant compared to those in the vertical plane (Campos Celador *et al.*, 2011; Kleinbach *et al.*, 1993). There is less agreement on the number of layers necessary to fully represent a tank, from anywhere between 3 to 200 (Rosen, 2001). However, any existence of layers is ignored by current linear optimisation models. Indeed, Steen *et al.* (2015) admitted that temperature dependent storage could not be modelled in the linear framework DER-CAM due to the nonlinearities imposed by temperature tracking. Instead, models suppose the storage medium is ideally stratified or perfectly mixed, with no vertical heat gradient. As a result, models only need to consider the energy charge and discharge and some percentage self discharge of the vessel.

Distribution networks are also necessarily simplified. Thermal networks are subject to more thermodynamic complexities than storage vessels, given that the energy carrier is not static. In fact, Li *et al.* (2016b) considered the distribution network as the heat storage medium for a district, which required mass continuity, temperature mixing, pressure losses, and flow rate limits to be modelled in detail. They found that they could improve on the overall economic efficiency of the district through detailed consideration of the distribution network, but were required to undertake nonlinear optimisation to achieve this.

Successful attempts have been made to improve on linearisation methods in linear energy system models. Where technology operating constraints are nonlinear, curves may be piecewise linearised. Through this, a curve can be approximated by several linear pieces. Wakui and Yokoyama (2014) applied piecewise linearisation to FC CHP technology efficiencies and HTPs when modelling a building energy system. Unlike a gas engine CHP, which has a higher rated efficiency, they argued that FC technologies have relatively higher efficiency at part-load conditions. Indeed, they found that a solid oxide FC CHP unit provided greater utility in optimal operation when piecewise linearised efficiencies were included. Evins *et al.* (2014) applied piecewise linearisation to heat pump efficiency in a building energy system, finding that the error in describing the curve reduced as the number of pieces increased. However, this error was only calculated ex-ante, which neglects the impact of the piecewise curve on the objective function value. The number of pieces was studied in greater detail by Bischi *et al.* (2014), who modelled several piecewise part-load efficiency curves for CHPs and for heat pump (HP) units, increasing the number of pieces from 5 to 20 in different

optimisation runs. They found that optimisation results did not vary, whether 10 or 20 pieces were used, suggesting that 10 pieces is sufficient to describe a technology characteristic curve.

Piecewise linearisation is certainly the most prevalent method of improving technology representation. It is used to describe technology efficiencies (Capuder and Mancarella, 2014; Fazlollahi *et al.*, 2014; Voll *et al.*, 2013; Voll *et al.*, 2015; Yokoyama *et al.*, 2014), technology costs (Gabrielli *et al.*, 2018; Majewski *et al.*, 2017), and energy distribution parameters such as pressure, temperature, and AC power flow (Cesena and Mancarella, 2018; Tanaka *et al.*, 2017). It is not the only method, however. Khir and Haouari (2015) iteratively solved linearised temperature and pressure in a heat network. But, a greater number of decision variables resulted in a prohibitive increase in solution time. Wakui *et al.* (2014) also modelled heat network temperature levels by considering heat to have lost usefulness after a set time period, on the grounds that temperature will have dropped. This could be used to model heat retention in the distribution, which acts as a potential storage system. Similarly, Fazlollahi *et al.* (2014) considered a storage vessel as a set of discrete fixed-volume temperature levels. Heat stored in a higher temperature level will reduce over time, filtering down into lower temperature levels as it loses usefulness.

## 2.2.2 Model determinacy

Linear optimisation is deterministic, that is, the optimal solution for a given set of input parameters will always be the same. However, the deterministic nature of the model has its limits; the optimum is only valid for the values assigned to the set of input parameters and the specific combination of those parameters in the objective function. Any adjustment to the parameters or objective is likely to change the value assigned to a decision variable. A study which analyses a single model configuration is thus ignorant of the robustness to uncertainty associated with their optimal technology investment portfolio.

### Uncertainty

Sensitivity or scenario analysis are the standard approaches to handling parameter uncertainty or updating the objective. In both approaches, several deterministic models are created, each with different values assigned to particular parameters. The difference between models is analysed, to compare the impact of parameter variation on the **objective function value (OFV)**. In sensitivity analysis, a single parameter is varied between models, thus assessing the sensitivity of the **OFV** to that parameter. If the set of runs does not vary significantly in **OFV**, then the system is not sensitive to the value assigned to a particular parameter. Wouters *et al.* (2015) used sensitivity analysis to consider the impact of energy prices, finding that the **OFV** increases with gas price increases, but not electricity price increases. However, the optimal distribution network changes most with electricity price changes.

Several parameters may be varied in scenario analysis. A modeller can compare various combinations of parameters that they deem plausible, although it is difficult to attribute impact on the **OFV** to any one parameter. Omu *et al.* (2015) considered scenarios in which the maximum capacity of technologies is varied. As such, they were able to simulate the possibility of a technology existing, or not, in the model. In doing so, they compared systems

with demand met by all-electric technologies, building-level technologies, or district-level technologies. Changing the scenario resulted in different system costs, CO<sub>2</sub> emissions, and air quality, although no attempt was made to include all possible technologies to see what the cost-optimal technology combination would be. Rather than just limit the available technologies for selection in planning optimisation, some studies use scenarios with fixed technology capacities so that only the operation schedule needs to be optimised, thus simplifying the problem at the expense of heuristically selecting capacities (Orehounig *et al.*, 2015; Wouters *et al.*, 2014).

Scenario and sensitivity analyses do not necessarily increase understanding of the robustness of a system. Although they can be useful to see how different parameter sets would influence a decision maker, they do not inform a decision maker on the risk of making a decision before uncertainty is realised, nor do they help mitigate this risk. Studies which seek to build uncertainty into the model use **robust optimisation (RO)** or **scenario optimisation (SO)**. In **RO**, the optimisation problem is reformulated to include uncertainty sets such that constraints are 'immunised' against infeasibility caused by the realisation of any value in a set (Li *et al.*, 2011). The degree to which a constraint is immunised is a function of a subjectively set level of 'conservatism'. Akbari *et al.* (2016) used **RO** to account for demand uncertainty in an illustrative district network. Here, demand is described by a 'box' uncertainty, which is akin to a uniform distribution. As conservatism increases, system cost increases. Concurrently, dependence on grid electricity and the possibility of unmet demand decreases. Zhou *et al.* (2018) similarly used a box uncertainty to describe wind, solar, price, and demand uncertainty when modelling an illustrative district over one day. They studied the participation of power-producing users in the national energy market and found that there is less interaction with the market as conservatism increases. There is also a greater contribution of conventional power generation to mitigate renewable power uncertainty.

In contrast to **RO**, which defines the outer limits of an uncertain parameter, **SO** defines distinct scenarios in which parameter uncertainty is realised. In a planning situation, decisions on technology capacity are made without knowing which of the future scenarios will be realised. The investment must be able to meet demand in any scenario, and minimise costs across them. More likely scenarios are given greater weighting, and consequently they have greater impact on the objective function. Pazouki and Haghifam (2016) sampled from uniformly distributed uncertainty sets with a 10% variance to generate scenarios with differing wind, energy price, and demand profiles. The introduction of uncertainty reduces dependence on **WTs** in favour of **CHPs**. Mavromatidis *et al.* (2018a) considered uncertain demand and available solar energy in a small Swiss district, finding that **NGB** and **PV** capacity increased to account for the uncertainty.

**SO** is advantageous, precisely because it does not require uncertainty to be strictly bounded. Additionally, conservatism can be applied disproportionately to the scenarios. Disproportionate conservatism is used to penalise the realisation of scenarios which would lead to particularly unfavourable events, even if the probability of this happening is low. This is known as applying risk-averse **SO**, which stands in contrast to the aforementioned risk-neutral **SO**. Mavromatidis *et al.* (2018c) applied various risk aversion techniques to

their previously risk-neutral **SO** model, finding that there is little agreement between them on the optimal investment capacity, but they all reduce the cost associated with the most unfavourable scenarios.

Regardless of whether uncertainty is accounted for in the objective function, all linear models remain deterministic. That is to say, for the same inputs, the output will be the same. As such, it is useful to test optimisation results when exposed to new information. Mavromatidis *et al.* (2018c) did not do this, and it is, therefore, difficult to understand the true benefit of having accounted for uncertainty, i.e. what would actually happen if the uncertainty is realised in a way that was not previously accounted for. Akbari *et al.* (2016) undertook an ex-post calculation of demand which would not be met by two sets of investment portfolios, one resulted from risk-unaware optimisation and the other from **RO**. When using high levels of conservatism in **RO**, unmet demand reduced by more than 90% compared to the risk-unaware model. Pazouki *et al.* (2014) constructed four different indicators associated with unmet electricity demand to analyse the performance of **SO** results when exposed to new operating conditions. **SO** resulted in the addition of **StoreEs** and a **CHP** in the investment portfolio, which led to a better performing system according to all four indicators.

## Objective

Typically, studies will consider monetary cost minimisation as the model objective in their study. In reality, there are likely to be competing objectives when designing a system. Thus, an investment portfolio needs to be robust to a change in objective, as well as to uncertainty. Omu *et al.* (2015) calculated the emitted CO<sub>2</sub> and health-related air pollutants following cost minimisation of several technological scenarios, showing that the cost-optimal solution would not necessarily have met other objectives, in this case, the minimisation of pollutant emissions. When the minimisation of CO<sub>2</sub> emissions is an objective of a study, it is possible to incorporate their impact directly into a cost minimisation objective function, by use of a weighting factor (Li *et al.*, 2016a) or by monetisation (Al Rafea *et al.*, 2017; Chen *et al.*, 2016; Christidis *et al.*, 2012; Ren *et al.*, 2010). The latter approach can incorporate the existing market prices of carbon in a model, which may help balance carbon emissions and monetary cost in an optimal solution. Although there are market prices for carbon emissions, such as those resulting from the European Union emissions trading scheme (Hirst, 2018), they are currently imposed on large-scale power producers and heavy industry, not on small district systems. Instead, some studies concentrate on how a limit on system carbon emissions may affect a system (Buoro *et al.*, 2013; Majewski *et al.*, 2017; Mavromatidis *et al.*, 2018c). Incrementally tightening a carbon emission limit leads to several solutions which describe a cost-carbon trade-off curve, also known as a 'Pareto front'. The Pareto front allows a designer to see how a reduction in carbon emissions could impact cost and therefore make a final investment decision independently of any one optimisation result.

Concurrently optimising multiple objectives, known as multi-objective optimisation, is not possible in a single **MILP** model without the aforementioned use of weighting factors or the production of pareto-fronts. Where studies have undertaken multi-objective optimisation, the **MILP** component of the model has been a part of a multi-stage optimisation. Fazlollahi

*et al.* (2014) 'enslaved' their linear model, which optimised the operation schedule for cost minimisation, with a 'master' evolutionary algorithm. The master optimisation attempted to co-minimise cost, carbon, and energy use based on limited knowledge of technology operation. On each iteration of the master optimisation, the slave optimisation would run to give a detailed understanding of optimal operation to feed back to the master. Multi-stage models such as this must then iterate for a set period of time, or until an optimal investment portfolio is converged upon. Ren *et al.* (2010) also considered a multi-stage model, where all levels were linear. The master model similarly optimised the investment portfolio, but was informed by two independent MILP slave models which were independently optimising for monetary cost or carbon emissions, respectively. Although the models studied by Ren *et al.* (2010) and Fazlollahi *et al.* (2014) offer a clearer understanding of the required trade-offs between cost and carbon, they were necessarily small. The iterative process of converging on a solution between the slave model(s) and the master can be time consuming. Thus, both studies considered their district as only a single node and limited the timeseries to seven TDs (Fazlollahi *et al.*, 2014) and three consecutive days (Ren *et al.*, 2010).

Research into balancing objectives tends to focus on carbon emissions, but other objectives are also studied. Indeed, robustness to uncertainty can be considered a risk-minimisation objective, when conducting RO or SO. Reddy (2017) included the minimisation of system power losses as part of the objective of their electrical district network, while Chen *et al.* (2016) and Zhang *et al.* (2015a) monetised the emission of pollutants other than CO<sub>2</sub> while also aiming to minimise cost. Energy use minimisation is also considered by Al Rafea *et al.* (2017) and Fazlollahi *et al.* (2014). However, the purpose of this is less clear, since the consumption of energy sources external to the district is always minimised, albeit following scaling by the monetary cost or carbon emissions associated with that energy source. Perhaps more important would be the minimisation of exergetic losses. In such a model, the conversion of more 'useful' energy (e.g. electricity) to less useful energy (e.g. low temperature heat) would be considered as a poor decision, due to the loss of exergy in the conversion process. Exergy analysis is nonlinear, however, as it depends on temperature tracking. It has thus currently only been considered by fully nonlinear models (e.g. Wang *et al.*, 2017).

Ultimately, the monetary cost of a system is the driving force in practical decision making. However, investment and operational costs may not be equally balanced in the way current planning studies choose to define the cost-optimal objective. Recognising this, Fazlollahi *et al.* (2014) studied the investment-operation cost Pareto front resulting from a cost-carbon multi-objective optimisation. Lower investment cost decisions incurred a higher operation cost, but lower carbon emissions. Consequently, the degree of reduction in carbon emissions depends on whether the decision maker is responsible for investment, operation, or both.

### 2.2.3 Model transparency

The studies reviewed in Section 2.1 frequently disagree on the validity of district systems as an alternative to conventional energy technologies. Carbon emissions could fall by 23% (Morvaj *et al.*, 2016), 44% (Li *et al.*, 2016a), or 50% (Omu *et al.*, 2013). On the other hand, they

could also more than double, albeit with a reduction of 74% in system cost, as reported by Jennings *et al.* (2014) and Mehleri *et al.* (2012). Cost savings are not universal either. Buoro *et al.* (2014) suggests only minor savings of a few percent in cost, while Haikarainen *et al.* (2014) would expect a slight increase.

Differences in results should be expected, since the case study district of each study is largely unique and the boundaries of the model change accordingly. The setting for each district will lead to distinct demand profiles, weather conditions, cost structures, and technology sets. Understanding these differences can only occur if input data can be inspected. Therefore, it is important to be clear on the data used to inform a model, and to communicate when such data is based on simulations, expert judgement, or external sources. However, in some studies (e.g. Carpaneto *et al.*, 2015; Good and Mancarella, 2017; Kotzur *et al.*, 2018a), parameter values are only partially published. When data is available for scrutiny, the choice of parameter values is often unclear or unsupported. For instance, Section 2.1.2 demonstrated a clear lack of agreement on the efficiency of various technologies; few studies gave a source for their chosen values.

In addition to lack of data transparency, the models themselves are rarely made publicly available. As noted by DeCarolis *et al.* (2012), if a model is not publicly available it can lead to concerns that there may be 1) hidden flaws or bugs in the source code or data; 2) subjective or value-based assumptions driving the results, and; 3) an obscured or absent effect of highly sensitive parameters in the published analysis. Additionally, unless it is a primary methodological contribution of the study, it is difficult to ascertain the complexity with which particular physical phenomena have been modelled. There is little incentive to make models openly available in the academic community, where *Energy Economics* is the only relevant journal requiring source code to be published alongside analysis (Pfenninger, 2017b). In the reviewed literature of district energy models, only DGOPT (Tanaka *et al.*, 2017) and ETEM-SG (Babonneau *et al.*, 2017) have made their source code publicly available. The two models offer two extremes of the open modelling movement. ETEM-SG is part of a suite of open source software provided by the ORDECSYS research group (<http://www.ordecsys.com>), while DGOPT is the initiative of an individual to make source code publicly available (<https://github.com/ikki407/DGOPT>).

Where models are closed, it is often due to their absence in any repository; the software developed for a study, or a set of studies, may not have been formalised as software. Although it may be possible to gain access by getting in contact with the author(s), DeCarolis *et al.* (2012) did not consider these models as publicly accessible since it creates a point of control that could be used to selectively or unwittingly deny access. Of those models which are published, only DER-CAM is specified as closed-source. Yet, anybody can verify the results of DER-CAM studies, as the packaged software is available for download (<https://building-microgrid.lbl.gov/projects/der-cam>), even if the source code is opaque.

Barriers to verification also stem from the software chosen to construct a model. All studies which divulge the optimisation software used to solve their models use IBM ILOG CPLEX (e.g. Garcia and Weisser, 2006; Zhou *et al.*, 2018). This is a proprietary piece of software for which a free academic licence is available. As well as being a linear optimiser, CPLEX

offers a user interface for the definition of a model. However, doing so restricts the model to being solved only by CPLEX. As such, these models cannot be verified or used by those outside the academic community without incurring a high cost. GAMS is another software used to construct models (e.g. Vahid-Pakdel *et al.*, 2017; Voll *et al.*, 2013), but it has no free academic tier. Even if models are shared, they cannot be considered publicly available if the user is required to pay a licence fee for the software needed to use and verify the models.

Recently, there have been increased calls for publicly available data and model source code, to enable the enhancement of the energy modelling community (Cao *et al.*, 2016; Morrison, 2018; Pfenninger, 2017b; Pfenninger *et al.*, 2017; Pfenninger *et al.*, 2018). To this end, European researchers formed the Open Energy Modelling Initiative in 2014 (<http://www.openmod-initiative.org/>). Since then, the initiative has collated over 30 models. Of these, only Calliope, Balmorel, Oemof, PyPSA and URBS could feasibly be used for district energy models, as these models incorporate the necessary constraints and inputs for multi-energy systems. In general, the drive towards greater openness has focussed primarily on national and international models. This is perhaps understandable, as these models have the potential to influence multinational policies such as the European Commission's Energy Roadmap 2050 (DeCarolis *et al.*, 2012). In contrast, district-scale models are primarily methodological, without any explicit intent on influencing policy as of yet. This should not preclude active involvement in the open modelling movement, however, since district-scale research could become increasingly more important in the decision-making process. Whether looking at a single development or designing policy on a local government level, models will prove more useful if decision makers can engage directly with them and understand the assumptions on which they are grounded (Pfenninger *et al.*, 2017).

## 2.3 Challenges addressed in this thesis

The previous section discussed key shortcomings in the current district energy optimisation paradigm: model simplification, determinacy, and transparency. Each has been addressed to some extent by recent studies in the field of MILP district energy optimisation, but they do not fully explore the impact of their methods on the decision-making process. As such, this thesis seeks to address four key challenges perceived to be critical for enabling practical applications of linear district energy system optimisation. By overcoming these challenges, a tool will be developed which is usable beyond academic and methodological analysis of district energy systems.

First, a modelling framework is developed to implement and test methodological enhancements in a transparent and validateable way. Chapter 3 introduces the open-source modelling framework that has been developed to address the three subsequent challenges, while remaining openly available to decision makers as a user-friendly tool. This availability is particularly novel as it allows practitioners to apply the research outputs from this thesis, with the opportunity to also extend functionality to suit their needs. Indeed, the modelling framework itself has been adapted from its original use as a national-scale energy system

tool, in order to accommodate district energy systems. Thus, this thesis offers a blueprint for extensibility, while itself offering transparent insights into district energy system design.

Second, model simplifications are explored in Chapter 4 through the application of piecewise linearisation. Unlike previous studies, piecewise linearisation is studied in greater detail in this thesis. This will entail the quantification of both ex-ante and ex-post error; the direct comparison between a nonlinear model, piecewise linear model, and rated efficiency model; and the possibility of optimising piecewise curve breakpoint allocation. Metaheuristic and MILP algorithms are applied to the same building-level illustrative case study model. A direct comparison such as this has not yet been undertaken in existing energy system models. Where piecewise linearisation is concerned, improvements are required to increase prevalence of its use in MILP models. This drives a detailed study into piecewise linearisation methods, beyond the direct comparison with fully nonlinear and rated efficiency cases.

Third, a three-step methodology is developed in Chapter 5 to incorporate uncertainty in a model. By clearly setting out the process by which a modeller can generate scenarios, prepare a tractable SO model, and quantify the impact of model results, it is possible to develop a practical tool. A key element of the method is that scenario generation is demand-driven. Unlike existing studies, which assume the distribution of uncertain parameters, a data-driven approach uses historical data to inform the generation of multi-dimensional probability density functions. This approach facilitates validation of the model and its practical use by decision makers who have access to data from existing, similar developments.

Fourth, the impact of competing objectives is examined in Chapter 6. Using the same case studies, carbon, cost and risk objectives are modelled both independently and together. Objectives are assigned to prospective decision makers, and their robustness to the realisation of uncertainty is quantified. In doing so, it is possible to compare the inadvertent impact of a decision across objectives. Additionally, the ability to incorporate multiple objectives into SO is explored. In particular, SO is enhanced to make most use of the risk aversion component, which is seldom used in existing district energy system optimisation models.



# Chapter 3

## Energy system optimisation

The development of Calliope, the modelling framework introduced in this chapter and used throughout this thesis, has been undertaken in collaboration with Stefan Pfenninger (ETH Zurich). Stefan originally developed the software for use in long term planning of national scale infrastructure. This thesis will focus on the author's contribution, comprising almost 500,000 lines of code in the Python programming language and two major releases which have been downloaded a total of 13,000 times. Parts of this chapter have been published as:

S. Pfenninger and B. Pickering (Sept. 2018). 'Calliope: A Multi-Scale Energy Systems Modelling Framework'. en. In: *The Journal of Open Source Software* 3.29, p. 825. DOI: [10.21105/joss.00825](https://doi.org/10.21105/joss.00825)

[Mixed integer linear programming \(MILP\)](#) is the primary method utilised within this research. This chapter begins by outlining the computational method by which [MILP](#) can be used to solve linear energy system optimisation problems, followed by the detailed description of the mathematical formulation of an energy model. The latter half of this chapter focusses on the open source modelling framework Calliope, which has been extended in this research to enable district-scale systems to be modelled and methodological enhancements to be examined. Calliope is capable of translating energy system models to a mathematical representation suitable for optimisation. Once optimisation is complete, the framework can package and visualise the data for clear dissemination. Furthermore, the framework operates under an open, permissively licensed Python tool-chain. As such, this chapter presents Calliope as a practical tool to address the challenge of model transparency.

The two sections of this chapter work as a pair to present energy system optimisation models. Figure 3.1 shows the workflow connecting the two. Calliope is used for the preparation and analysis of models. [MILP](#) is an algorithm for the optimisation of these problems. Due to the established nature of [MILP](#), energy system modellers do not attempt to improve on the algorithmic approach of their models. Indeed, this separation of framework and algorithm is standard within the district energy modelling community, where linear models are concerned. Models such as DENO, HOMER, URBS, and DER-CAM are more accurately defined as *frameworks*, which in turn depend on linear optimisers, also known as *solvers*, such as IGM ILOG CPLEX (IBM Corp., 2016) and Gurobi (Gurobi, 2018).

Existing frameworks provide similar functionality. Indeed, this thesis began by using DENO, written natively in IBM ILOG CPLEX, before moving to the development of Calliope.

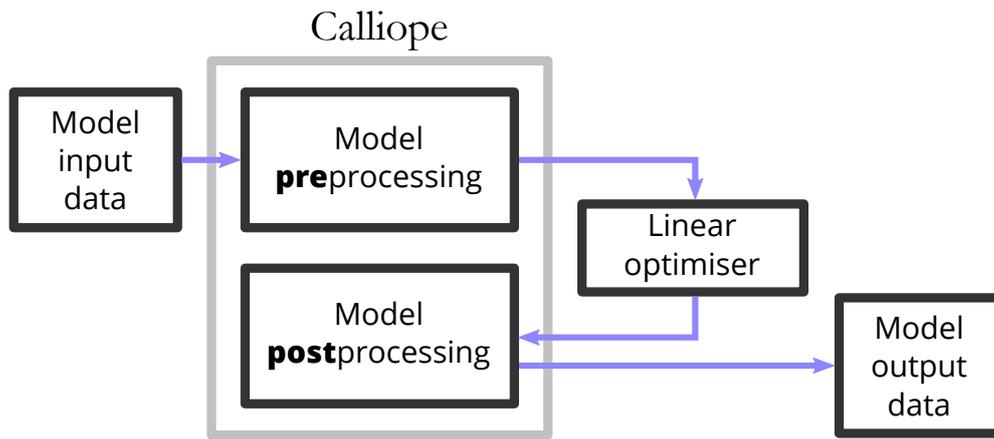


Figure 3.1 Workflow connecting the Calliope model framework and the linear optimiser, otherwise known as the linear solver.

Differences between the available frameworks lie in accessibility and usability. Frameworks which boast a user interface, such as DER-CAM and GAMS, are often inaccessible, through restrictive licencing or inexistent packaging of the work. Although openness is becoming more prevalent, as evidenced by the array of software collated by the Open Energy Modelling Initiative (<https://openmod-initiative.org>), openly available frameworks, such as URBS and Oemof, do not prioritise usability. Uniquely, Calliope is designed for use by a wider audience and is openly available for its capabilities to be extended. As the aim of this thesis is to provide a practical tool for the use of mathematical optimisation by decision makers, Calliope proves to be an ideal candidate. Usability, alongside the more generic formulation of an energy system model, will be discussed in greater detail in Section 3.2.

### 3.1 Mixed integer linear optimisation

**Linear programming (LP)**, sometimes referred to as linear optimisation, is an extensively used method for solving a given mathematical function. This function, known as the *objective function* has a number of *decision variables* for which a modeller aims to find suitable values. Depending on the aim of the problem, the value of the objective function may need to be maximised (e.g. system revenue) or minimised (e.g. system cost). The variables can take on any continuous number, but are constrained by additional functions which bound the feasible space of solutions. The general formulation of the objective function and constraints applied in an LP model is given in Equation 3.1. As the name suggests, nonlinearities cannot be represented in an LP model. Consequently, this precludes the multiplication of two decision variables in a constraint or in the objective function.

$$\begin{array}{ll}
\text{Minimise or maximise} & c_1x_1 + c_2x_2 + \dots + c_nx_n \\
\text{subject to} & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n (\leq, =, \text{ or } \geq) b_1 \\
& a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n (\leq, =, \text{ or } \geq) b_2 \\
& \dots \\
& a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n (\leq, =, \text{ or } \geq) b_m \\
& x_j \geq 0 \quad \forall j = 1, \dots, n.
\end{array}$$

Where  $x_1, \dots, x_n$  is the set of decision variables;  $c_1, \dots, c_n$  are the objective coefficients;  $b_1, \dots, b_m$  are the constraint limits, and;  $a_{11}, \dots, a_{ij}, \dots, a_{mn}$  are the constraint coefficients, which denote how much of a requirement  $i$  is satisfied by the decision  $j$  (Smith and Taskin, 2008).

Figure 3.2 shows a bounded region for two decision variables,  $x$  and  $y$ . The objective function,  $z$ , to be maximised is  $2x + 3y$ . Without any constraints (Figure 3.2a), this is an unbounded problem; the solution is infinitely large, as  $x$  and  $y$  can both continue to infinity. Bounding the problem, using the five lines given in Figure 3.2b, provides us with a maximised objective function value of 10. It is clear that within the solution space, known as the *convex hull*, decision variables acquire their optimal values at a vertex. Capitalising on vertex optimality, the popular *simplex* algorithm inspects only vertex values, sequentially ‘climbing’ the vertices of an  $N$ -dimensional convex hull to reach the optimal solution (Figure 3.3). It is important to understand this mechanism, as constraints on decision variables can also lead to model infeasibility. Returning to the two-dimensional example, Figure 3.2c shows how two conflicting constraints on  $y$  means that there is no possible value that  $y$  can take.

Some variables within a model may not be continuous. Indeed, it is realistic to model some physical phenomena discretely. For instance, a technology is switched on (1) or it is not (0). Technologies will also be sold by manufacturers in discrete sizes, leading to any integer number of technologies to be purchased for a system. The incorporation of these decision variables requires integer linear programming or, because there still exist continuous variables in the formulation, *mixed integer linear programming* (MILP). Unlike an LP model, the solution to the problem might not lie on a vertex. If  $y$  can only be an integer value, the aforementioned two-dimensional example now has an optimal solution of nine, as depicted in Figure 3.2d. This is less optimal than the continuous solution, but more physically representative.

Reaching an optimal solution in a MILP model is computationally harder than in its continuous counterpart. The reason for this becomes clear when breaking down the solution method. First, a MILP algorithm will relax constraints to allow all variables to be continuous. If the solution to the linear relaxation does not already satisfy the integer constraints, the closest possible solution must be found by partitioning the convex hull. This convergence of the solution on that given by the linear relaxation is the cause of increases in solution time, as the initial linear optimisation will take as long to solve as it would do if all variables were linear (i.e. LP). The simplest technique for converging on the linear relaxation is to iterate over every possible integer solution. However, in practice, this is infeasible as the solution time

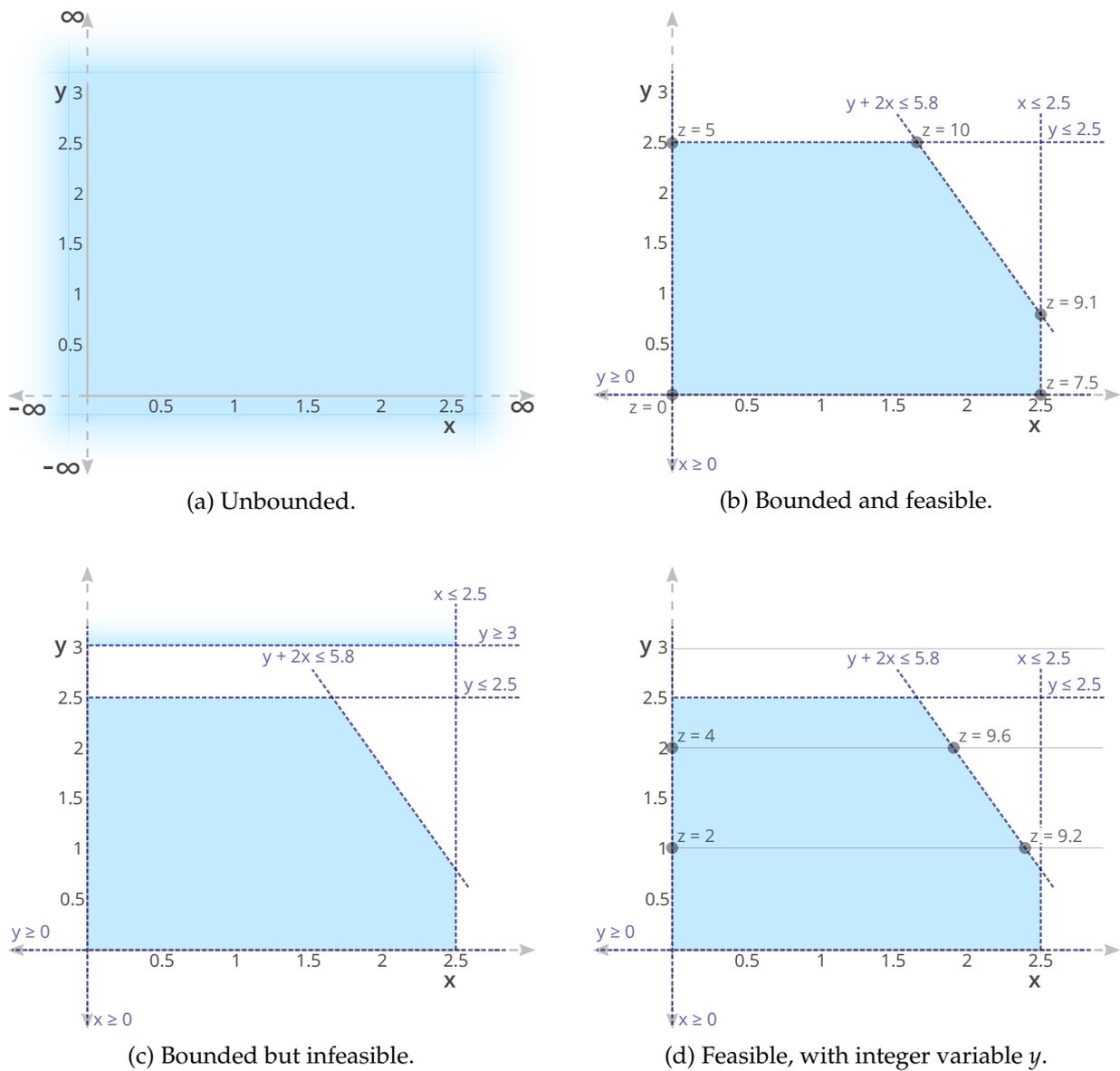


Figure 3.2 Bounded convex hull describing a two-dimensional LP problem. a), b), and c) are respectively unbounded, feasible, or infeasible given continuous variables  $x$  and  $y$ . d) is feasible considering the requirement for  $y$  to be an integer value. Objective function value  $z$  is given at feasible vertices.

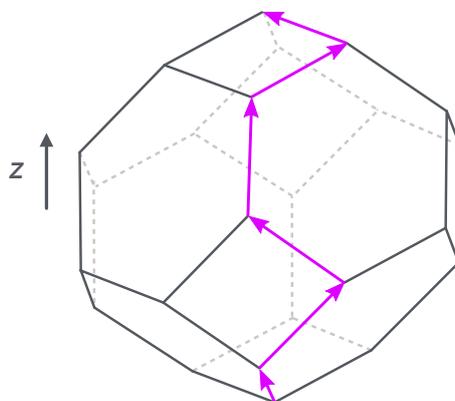


Figure 3.3 Graphical representation of an  $N$ -dimensional LP problem being solved by the simplex algorithm to maximise the objective function  $z$ . Figure modified from (Moore and Mertens, 2011).

increases exponentially with the number of variables. If a problem with only binary decision variables has  $n$  variables, then it will take  $2^n$  iterations to inspect all possible solutions. If  $n = 60$ , a high performance computer which can process 10 trillion operations per second will take 80 days to terminate. If  $n = 70$ , this increases to 262 years (Smith and Taskin, 2008).

Popular methods for converging on the linear relaxation of a MILP problem in a reasonable time are the ‘branch and bound’ and ‘cutting plane’ techniques. Both methods are well established, having been first introduced over 50 years ago (Gomory, 1960; Land and Doig, 1960). The branch and bound technique iteratively segments the convex hull, creating smaller hulls in which to search for an integer solution. Each time segmentation occurs, a new branch is created on the solution tree. This branch is searched until it is certain that no better solution can be found. The best solution of all branches is then taken to be the global optimum. The cutting plane works in reverse, by applying additional inequalities to remove possible vertex solutions from the linear relaxation without removing possible integer solutions. This continues until reaching the limit whereby the integer solution is on a vertex described by the cutting planes. Cutting planes are usually used in tandem with branch and bound, as part of the ‘branch and cut’ technique (Mitchell, 2002).

The branch and cut technique can lead to rapid convergence, particularly as it can be parallelised across multiple cores. However, this convergence can be slow when close to the solution of the linear relaxation. A common approach to mitigate this is to implement an allowable ‘gap’ which will end convergence early, provided the solution is within a certain percentage of the linear relaxation. Linear optimisers default the gap to 0.01%, but district energy studies tend to update this, allowing gaps of 0.1% - 1% (Buoro *et al.*, 2013; Gabrielli *et al.*, 2018; Omu *et al.*, 2013). In addition, most linear optimisers extend the simplex algorithm by applying heuristic methods to reduce the problem prior to applying simplex methods.

Solver choice also impacts solution time, which can be seen when the open-source optimisers GLPK (*GNU Linear Programming Kit, Version 6.1 2018*) and CBC (Forrest *et al.*, 2018) are compared with the two most applied commercial optimisers, CPLEX (IBM Corp., 2016) and Gurobi (Gurobi, 2018). A test model, included four interconnected nodes and nine technologies, was optimised over 8760 timesteps. Table 3.1 shows the several orders of magnitude increase in CPU time required by open-source solvers; the GLPK solver was manually stopped after five hours without reaching a solution. The problem began with 586,996 decision variables (continuous: 578,233, binary: 2, integer: 8,761) subject to 788,502 constraints. In the heuristic presolve phase, CPLEX reduced this to 249,377 decision variables subject to 258,122 constraints. Variables which can be discounted during heuristic presolving are those for which there is clearly only one value, such as setting a variable to a fixed value when defining the model. GLPK also undertakes a presolve phase, but it is less thorough, leaving 327,628 variables and 537,255 constraints. Both the simplex algorithm and the integer problem convergence are more powerful in CPLEX and Gurobi, resulting in a model which takes at least two orders of magnitude less time to solve than GLPK. However, CBC, another open-source solver, performs significantly better in comparison to its commercial counterparts, Gurobi and CPLEX. In fact, it solves the linear relaxation problem in 45 seconds, compared to 137 seconds for CPLEX, but converging on the integer solution takes 10 times

longer. Parallelisation across four threads reduces solution time, but is not available for the open-source solvers. In this particular problem, parallelisation does not particularly improve the best possible solution time, realised by Gurobi.

Table 3.1 MILP model solution time when run using different linear optimisers, showing the distinctly greater time to solution for open-source optimisers.

Solver	GLPK	CBC	Gurobi		CPLEX	
			1 thread	4 threads	1 thread	4 threads
<b>Solution time</b>	>5hrs	0:52:13	0:03:21	0:03:08	0:05:56	0:03:26

### Side note on tractability of optimisation models

Many of the methods researched in district energy optimisation, this thesis included, are driven by a requirement for tractability. In formal terms, tractability might be considered as the necessary condition for the problem to be computed in polynomial time, i.e. within complexity class  $P$  (Meurant, 2014, p.4). Thus  $NP - complete$  and  $NP - Hard$  (where  $P \neq NP$ ) would be considered intractable. This classic definition is somewhat questionable, as it is clearly more practical to solve a problem in the order of  $2^{0.1N}$  (exponential) than in the order of  $N^{20}$  (polynomial) for a large  $N$  (Rotman, 2003). Additionally, a problem that has complexity in the order of  $N^2$  becomes impractical to solve if the number of computations,  $N$ , is large (e.g.  $10^{10}$ ). Goderbauer *et al.* (2018) recently formalised the complexity of Mixed integer *nonlinear* optimisation energy system problems as  $NP - Hard$  but did not discuss MILP problems. In fact, model complexity is rarely formalised. Instead, tractability and practicality are interchanged. As such, a problem is considered to be tractable if it reaches an optimal (and feasible) solution in a ‘reasonable’ amount of time. What is reasonable depends on the modeller and the computational power at their disposal. Many studies wish to realise solutions to models in  $O(100s)$ . In this thesis, tractability is defined relative to available computing capacity and varies from  $O(1,000s)$  to  $O(100,000s)$ .

## 3.2 Calliope: an open source modelling framework

Calliope was the Greek muse of epic poetry. It is now also an open-source Python-based energy modelling framework. Initially released in 2013, the aims of the model were:

- to be easy to read;
- to allow the modelling of any mix of technologies, by only loosely defining components;
- to be highly resolvable, both spatially and temporally;
- to enable running models on high-performance computing (HPC) clusters;
- to be dependent only on state-of-the-art Python third party libraries, and;

- to be freely available under the Apache 2.0 licence.

The original purpose of the framework was to examine national-scale energy system models. The software has since been used to address research questions surrounding energy policy in the UK, Switzerland, and South Africa (Díaz Redondo and van Vliet, 2015; Pfenninger and Keirstead, 2015a; Pfenninger and Keirstead, 2015b). The aims of the software are theoretically scale independent, encompassing the formulation of any energy system model. In reality, when using the framework to prepare district scale models, the lack of focus on smaller scale, multi-energy systems led to inadequacies in the model. In this research, these issues have been addressed through two major releases of Calliope, namely:

- Calliope can now model MILP problems, not just LP ones;
- multi-energy technologies can be modelled, with linked and/or unlinked energy carrier flow;
- technologies can export energy outside the system, recognising the existence of districts as subsystems of national-scale systems;
- revenue generation is possible, through the introduction of negative costs;
- data visualisation is more accessible to users with low Python programming literacy;
- all model parameters and sets are more verbose and are aligned with energy sector terminology, to improve clarity of model definitions and results;
- models can be optimised for the operation schedule alone, using a rolling horizon approach, to test the validity of an investment portfolio when exposed to ‘real’ operating conditions;
- costs can vary in time, to simulate price fluctuations, and;
- models with high dimensionality are prepared for optimisation in a much shorter time period, by matrix densification.

The mathematical formulation of Calliope version 0.6 is detailed in this section, with particular focus on district energy system constraints and processes. Further details on its functionality are available in the project documentation (<https://calliope.readthedocs.io/en/stable/>).

### 3.2.1 Model formulation

A model of an energy system requires a user to define technologies, nodes, links, and energy carriers. Technologies sit within nodes and can interact with other technologies within a node, with other nodes via links, or with other entities outside the system boundary. Figure 3.4 illustrates this for a two-node district. The energy balance of each energy carrier present in a node must be satisfied within that node.

Supply and demand ‘technologies’ allow energy flow from and to other energy systems outside the district boundary, respectively. Only energy carriers within the system are explicitly defined, with carriers outside the system known as a ‘resource’. This may refer to a positive resource that can be supplied, e.g. natural gas, grid electricity, or solar energy, or to a negative resource that is demanded at a node, e.g. space heating, hot water, or electricity. Supply technologies consume a resource and produce an energy carrier, while demand technologies consume a carrier to produce a resource. Transmission technologies allow energy carrier flow between nodes within the system, by consuming it at one node and producing it at another. Conversion and storage technologies also interact only with system energy carriers, but this interaction can only take place within a single node.

### Decision variables, parameters, and sets

The purpose of optimisation is to calculate the values of *decision variables* which will minimise (or maximise) the objective function value. In an energy system model, these decision variables include technology capacity, energy production and consumption per timestep, and incurred costs. All decision variables in Calliope are given in Table 3.2. Each is indexed over a selection of dimensions, defined by the system sets and subsets (Table 3.3). Only the cost decision variable is required in the objective function; all others can be considered as ‘auxiliary’, since they impact cost but are not present in the objective function. Auxiliary variables are also usually the values in which a decision maker is interested and are consequently important to track independently from the total system cost.

Table 3.2 Decision variables defined in Calliope.

Decision variable	Calliope syntax	Chapter syntax	Domain	Dimensions
Energy capacity	energy_cap	$\hat{E}$	$[0, \infty)$	nodes, techs
Energy carrier production	carrier_prod	$E^+$	$[0, \infty)$	nodes, techs, carriers, timesteps
Energy carrier consumption	carrier_con	$E^-$	$(-\infty, 0]$	nodes, techs, carriers, timesteps
System cost	cost	<b>cost</b>	$(-\infty, \infty)$	nodes, techs, costs
Resource collection area	resource_area	$\hat{R}^{area}$	$[0, \infty)$	nodes, techs
Stored energy capacity	storage_cap	$\hat{S}$	$[0, \infty)$	nodes, techs
Stored energy	storage	<b>S</b>	$[0, \infty)$	nodes, techs, timesteps
Consumed resource <sup>a</sup>	resource_con	$R^-$	$[0, \infty)$	nodes, techs, timesteps
Resource capacity <sup>a</sup>	resource_cap	$\hat{R}$	$[0, \infty)$	nodes, techs, timesteps
Energy carrier export	carrier_export	$E^{ex}$	$[0, \infty)$	nodes, techs, carriers, timesteps
Operational cost	cost_var	<b>cost<sup>op</sup></b>	$(-\infty, \infty)$	nodes, techs, costs, timesteps
Investment cost	cost_investment	<b>cost<sup>inv</sup></b>	$(-\infty, \infty)$	nodes, techs, costs
Technology purchase switch	purchased	<b>P</b>	$\{0, 1\}$	nodes, techs
Purchased technologies <sup>a</sup>	units	$\hat{U}$	$\{0, 1, \dots, \infty\}$	nodes, techs
Technology unit dispatch	operating_units	<b>U</b>	$\{0, 1, \dots\}$	nodes, techs, timesteps
Unmet demand slack	unmet_demand	<b>slack<sup>+</sup></b>	$[0, \infty)$	nodes, carriers, timesteps
Unused supply slack	unused_supply	<b>slack<sup>-</sup></b>	$(-\infty, 0]$	node, carriers, timesteps

<sup>a</sup> Although defined here, these decision variables are not relevant to methods described in this thesis and are therefore not referred to hereafter.

Table 3.3 Sets and subsets defined in Calliope.

Sets	
$n \in nodes(N)$	Geographic nodes in the system
$(n, n_r) \in links(L)$	Node pairs, defining transmission links
$x \in techs(X)$	Available technologies (e.g. 'CHP')
$t \in timesteps(T)$	Operational timesteps (e.g. '2017-01-01 00:00:00')
$c \in carriers(C)$	Energy carriers (e.g. 'electricity', 'heat')
$cost \in costs$	Cost classes (e.g. 'monetary', 'carbon')
Subsets	
<i>store</i>	Storage technologies
<i>conv</i>	Conversion technologies
<i>supply</i>	Supply technologies
<i>dem</i>	Demand technologies
<i>trans</i>	Transmission technologies
<i>ex</i>	Technologies which can export a carrier
<i>prod</i>	Technologies which can produce a carrier
<i>con</i>	Technologies which can consume a carrier
<i>area</i>	Technologies which are constrained by physical area

Parameters are required to formulate the system constraints and the objective function. They include maximum capacity, available resource, and technology capacity/operation costs. Each technology subset can be constrained by a different selection of parameters. Therefore, parameters are listed in the following sections on a per-technology basis. However, only those parameters relevant to this thesis are given; a more detailed listing of parameters is available in the online documentation (<https://calliope.readthedocs.io/en/stable/>). In constraint definition, parameters are distinguished from decision variables by the latter being given in **bold** text.

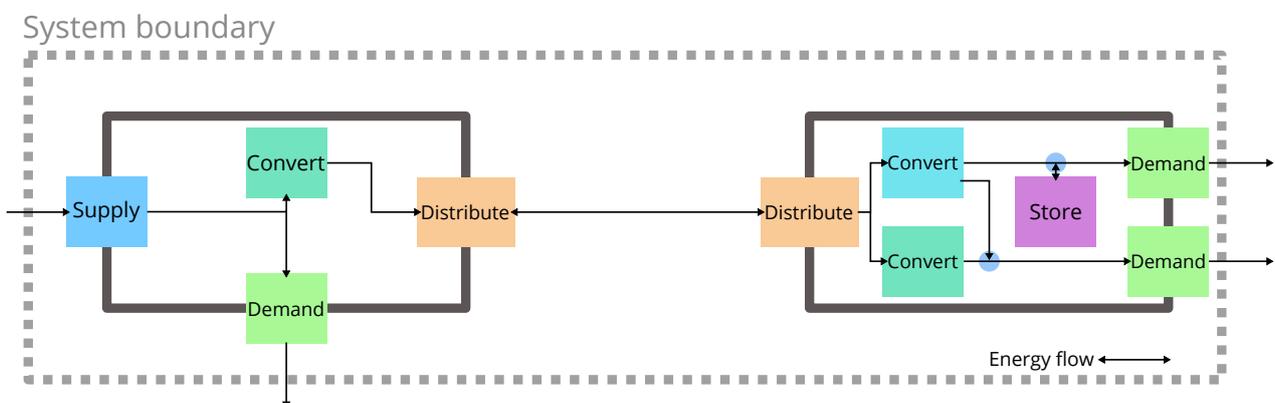


Figure 3.4 Illustrative two node district, with eight distinct technologies and four energy carriers.

## Constraints

At a technology level, constraints are built over smaller subsets of the system. For instance, in the example given in Figure 3.4, there is only a storage technology at node 2, consuming only one of the four energy carriers in the system. Thus storage constraints will be built over the storage ‘node::technology::carrier’ subset, which avoids the creation of superfluous constraints for this storage technology at node 1 or for any other energy carrier. To this end, the sparse matrix defining all possible node, technology, and carrier combinations (64 in the illustrative example) is compacted to a dense matrix of only the relevant node, technology, and carrier combinations (11 in the illustrative example). Hereafter, each constraint can be assumed to be built over the relevant dense subset.

Some constraints are technology agnostic:

- Maximum energy capacity

$$\hat{\mathbf{E}}_{n,x} \begin{cases} \leq \hat{E}_{n,x}^{max} \times \mathbf{P}_{n,x}, & \text{if } cost_{n,x}^P \geq 0 \\ = \hat{E}_{n,x}^{eq}, & \text{if } \hat{E}_{n,x}^{eq} \neq \text{None} \\ \leq \hat{E}_{n,x}^{max}, & \text{otherwise} \end{cases} \quad (3.1)$$

- Maximum carrier export

$$\mathbf{E}_{n,x,c}^{ex} \leq \mathbf{E}_{n,x,c}^+ \quad \forall (n, x, c) \in (N, X, C)_{ex} \quad (3.2)$$

Table 3.4 Valid parameters for constraining a supply technology.

Parameter	Calliope Syntax	Thesis syntax	Default	Domain
Verbose syntax				
Maximum installed energy capacity	energy_cap_max	$\hat{E}^{max}$	$\infty$	$[0, \infty)$ kW
Energy efficiency	energy_eff	$\eta^E$	1.0	$[0, \infty)$
Require resource consumption	force_resource	$R^{force}$	0	$\{0, 1\}$
Plant parasitic efficiency	parasitic_eff	$\eta^P$	1.0	$[0, \infty)$
Available resource	resource	R	0	$[0, \infty)$ kWh/m <sup>2</sup>
Maximum installed collector area	resource_area_max	$R^{area,max}$	$\infty$	$[0, \infty)$ m <sup>2</sup>
Collector area efficiency	resource_area_per_energy_cap	$\eta^{area}$	1.0	$[0, \infty)$
Resource scale	resource_scale	$R^{scale}$	1.0	$[0, \infty)$

**Supply** Energy supply technologies are those which consume a resource from outside the system boundary and produce an energy carrier within the system (Figure 3.5, Table 3.4). Specific constraints to apply are:

- Available resource

$$\frac{\mathbf{E}_{n,x,t,c}^+}{\eta_{n,x,t}^E \times \eta_{n,x,t}^P} \begin{cases} = R_{n,x,t}, & \text{if } R_{n,x}^{force} = 0 \\ \leq R_{n,x,t}, & \text{if } R_{n,x}^{force} = 0 \end{cases} \quad \forall (n, x, t, c) \in (N, X, T, C)_{supply} \quad (3.3)$$

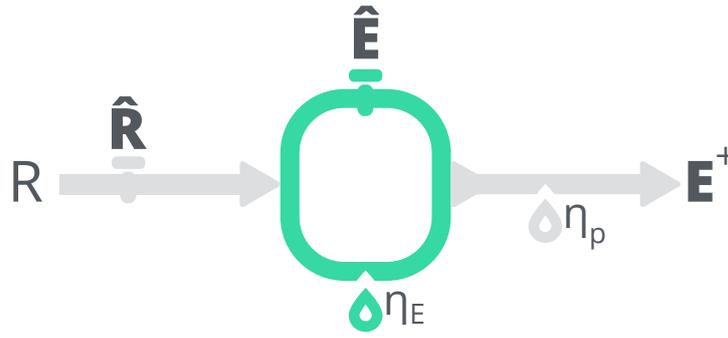


Figure 3.5 Calliope supply technology.

Where  $R^{force}$  is a binary parameter dictating whether a technology is dispatchable (e.g. diesel fuel) or not (e.g. un-curtailed [solar photovoltaic panel \(PV\)](#)).

- Maximum carrier production

$$\mathbf{E}_{n,x,t,c}^+ \leq \hat{\mathbf{E}}_{n,x} \times \eta_{n,x,t}^p \quad \forall (n, x, t, c) \in (N, X, T, C)_{prod} \quad (3.4)$$

- Available resource per unit area

$$\frac{\mathbf{E}_{n,x,t,c}^+}{\eta_{n,x,t}^E} \begin{cases} = R_{n,x,t} \times \hat{\mathbf{R}}_{n,x}^{area}, & \text{if } R_{n,x}^{force} = 0 \\ \leq R_{n,x,t} \times \hat{\mathbf{R}}_{n,x}^{area}, & \text{if } R_{n,x}^{force} = 1 \end{cases} \quad \forall (n, x, t, c) \in (N, X, T, C)_{supply,area} \quad (3.5)$$

- Resource area capacity

$$\hat{\mathbf{R}}_{n,x} \begin{cases} \leq \hat{\mathbf{R}}_{n,x}^{area,max} \times \mathbf{P}_{n,x}, & \text{if } cost_{n,x}^P \geq 0 \\ \leq \hat{\mathbf{R}}_{n,x}^{area,max}, & \text{otherwise} \end{cases} \quad \forall (n, x) \in (N, X)_{area} \quad (3.6)$$

- Area use to energy capacity linking

$$\mathbf{R}_{n,x}^{area} = \hat{\mathbf{E}}_{n,x} \times \eta_{n,x}^{area} \quad \forall (n, x) \in (N, X)_{area} \quad (3.7)$$

Table 3.5 Valid parameters for constraining a conversion technology.

Parameter	Calliope Syntax	Thesis syntax	Default	Domain
<i>Verbose syntax</i>				
Maximum installed energy capacity	energy_cap_max	$\hat{\mathbf{E}}^{max}$	$\infty$	$[0, \infty]$ kW
Energy efficiency	energy_eff	$\eta^E$	1.0	$[0, \infty)$
Technology lifetime	lifetime	L	0	$[0, \infty)$ years

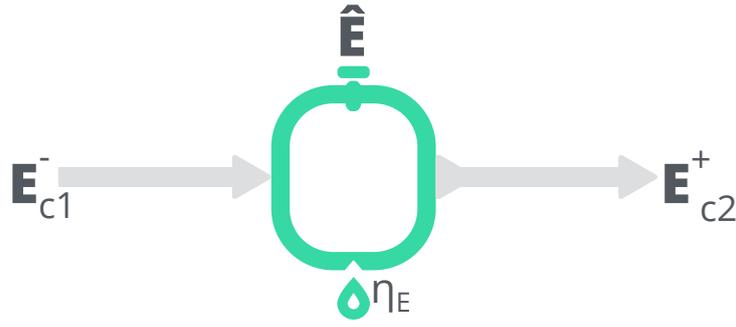


Figure 3.6 Calliope conversion technology.

**Conversion** Energy conversion technologies are those which consume an energy carrier from within a node and produce an energy carrier within the same node (Figure 3.6, Table 3.5). Specific constraints to apply are:

- Carrier conversion

$$-1 * \mathbf{E}_{n,x,t,c_{con}}^- \times \eta_{n,x,t} = \mathbf{E}_{n,x,t,c_{prod}}^+ \quad \forall (n, x, t) \in (N, X, T)_{conv} \quad (3.8)$$

- Maximum carrier production

$$\hat{\mathbf{E}}_{n,x,t,c}^+ \leq \mathbf{E}_{n,x} \quad \forall (n, x, t, c) \in (N, X, T, C)_{prod} \quad (3.9)$$

Table 3.6 Valid parameters for constraining a multi-energy conversion technology.

Parameter	Calliope Syntax	Thesis syntax	Default	Domain
<i>Verbose syntax</i>	<i>Calliope Syntax</i>	<i>Thesis syntax</i>		
Energy carrier conversion ratios	carrier_ratios	ratio	1	$[0, \infty)$
Maximum installed energy capacity	energy_cap_max	$\hat{E}^{max}$	$\infty$	$[0, \infty]$ kW
Energy efficiency	energy_eff	$\eta^E$	1.0	$[0, \infty)$
Technology lifetime	lifetime	L	0	$[0, \infty)$ years

**Multi-energy Conversion** Multi-energy conversion technologies are those which consume a number of energy carriers from within a node and produce a number of energy carriers within the same node (Figure 3.7, Table 3.6). Two streams of energy production or consumption may exist. Separate production/consumption streams ( $E_1^+$  and  $E_2^+$  in Figure 3.7) are linked, such that their sum in any timestep is linked by a fixed ratio. For example, the heat and power output of a **combined heat and power plant (CHP)** is linked by the **heat-to-power ratio (HTP)** (Figure 3.8a). Within a production/consumption stream, multiple carriers may be defined (e.g.  $E_{1,1}^+$  and  $E_{1,2}^+$  in Figure 3.7) that are not linked, but the sum of their production must equal the stream production (e.g.  $E_1^+$ ). This allows an **air source heat pump (AHP)** to be defined with the choice of heat or cooling output (Figure 3.8b). The **combined cooling,**

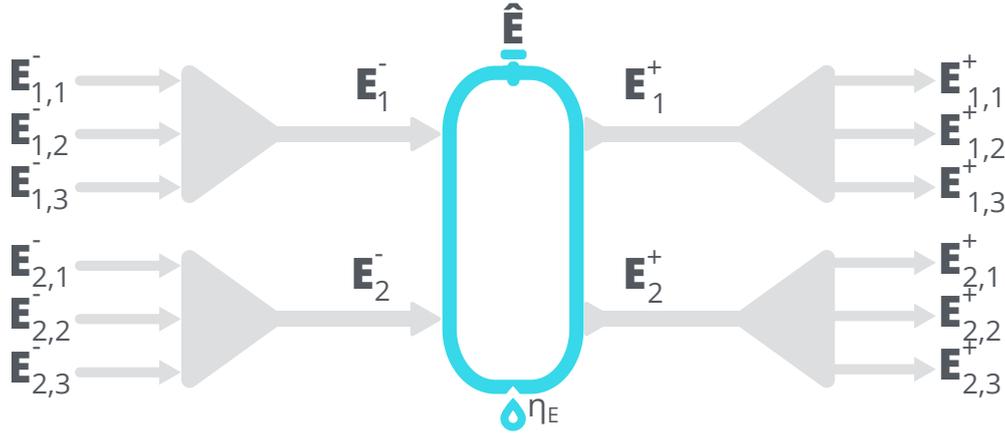


Figure 3.7 Calliope multi-energy conversion technology.

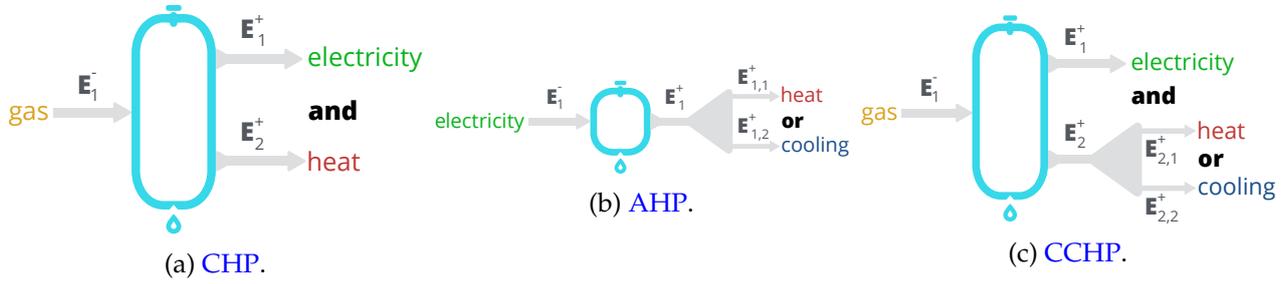


Figure 3.8 Representation of various multi-energy conversion technologies in Calliope, which may have energy carriers linked or unlinked.

heat and power plant (CCHP) in Figure 3.8c shows both linked streams and unlinked carriers within a stream. A ‘primary carrier’ is defined for each multi-energy conversion technology, to which Equation 3.9 is applied to constrain the technology maximum capacity. In addition, the following constraints are required:

- Carrier conversion

$$-1 * \sum_{c \in \text{stream}_{\text{con},1}} \frac{E_{n,x,t,c}^-}{\text{ratio}_{n,x,c,\text{con}_1}} \times \eta_{n,x,t} = \sum_{c \in \text{stream}_{\text{prod},1}} \frac{E_{n,x,t,c}^+}{\text{ratio}_{n,x,c,\text{prod}_1}} \quad (3.10)$$

$$\forall (n, x, t) \in (N, X, T)_{\text{conv}}$$

- Linking inter-stream production

$$\sum_{c \in \text{stream}_{\text{prod},1}} \frac{E_{n,x,t,c}^+}{\text{ratio}_{n,x,c,\text{prod}_1}} = \sum_{c \in \text{stream}_{\text{prod},2}} \frac{E_{n,x,t,c}^+}{\text{ratio}_{n,x,c,\text{prod}_2}} \quad \forall (n, x, t) \in (N, X, T)_{\text{conv}} \quad (3.11)$$

$$\sum_{c \in \text{stream}_{\text{con},1}} \frac{E_{n,x,t,c}^-}{\text{ratio}_{n,x,c,\text{con}_1}} = \sum_{c \in \text{stream}_{\text{con},2}} \frac{E_{n,x,t,c}^-}{\text{ratio}_{n,x,c,\text{con}_2}} \quad \forall (n, x, t) \in (N, X, T)_{\text{conv}} \quad (3.12)$$

Table 3.7 Valid parameters for constraining a storage technology.

Parameter	Calliope Syntax	Thesis syntax	Default	Domain
<i>Verbose syntax</i>				
Carrier ratios	carrier_ratios	ratio	1	$[0, \infty)$
Charge rate	charge_rate	$\eta^{charge}$	0	$[0, \infty)/\text{hour}$
Maximum installed energy capacity	energy_cap_max	$\hat{E}^{max}$	$\infty$	$[0, \infty]\text{kW}$
Energy efficiency	energy_eff	$\eta^E$	1.0	$[0, \infty)$
Maximum storage capacity	storage_cap_max	$\hat{S}^{max}$	$\infty$	$[0, \infty]\text{kWh}$
Initial storage level	storage_initial	$S^{init}$	0	$[0, \infty)\text{kWh}$
Storage loss rate	storage_loss	$loss^S$	0	$[0, 1]/\text{hour}$

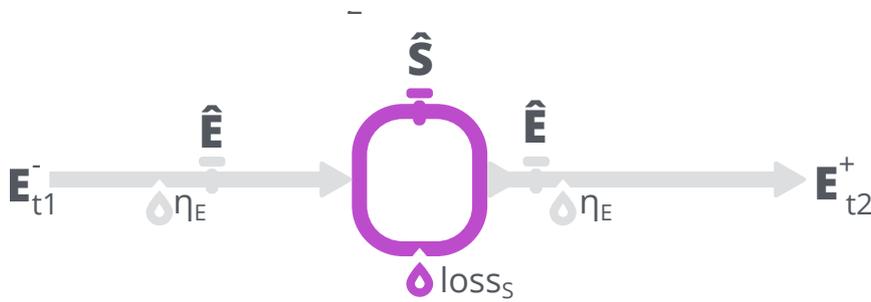


Figure 3.9 Calliope storage technology.

**Storage** Energy storage technologies are those which consume an energy carrier from within a node and produce the same energy carrier within the same node, but at a different timestep (Figure 3.9, Table 3.7). Specific constraints to apply are:

- Inter-temporal energy balance

$$\mathbf{S}_{n,x,t} = \begin{cases} \mathbf{S}_{n,x,t-1} \times (1 - loss_{n,x,t}^S)^{res_t} - \mathbf{E}_{n,x,t,c}^- \times \eta_{n,x,t}^E - \frac{\mathbf{E}_{n,x,t,c}^+}{\eta_{n,x,t}^E}, & \text{if } t > 0 \\ S_{n,x}^{init} - \mathbf{E}_{n,x,t,c}^- \times \eta_{n,x,t}^E - \frac{\mathbf{E}_{n,x,t,c}^+}{\eta_{n,x,t}^E}, & \text{if } t = 0 \end{cases} \quad (3.13)$$

$$\forall (n, x, t) \in (N, X, T)_{store}$$

Where *res* is the timestep resolution.

- Linking storage and energy capacity, where the latter is synonymous with the charge/discharge rate

$$\hat{\mathbf{E}}_{n,x} \leq \hat{\mathbf{S}}_{n,x} \times \eta_{n,x}^{charge} \quad \forall (n, x) \in (N, X)_{store} \quad (3.14)$$

- Stored energy capacity

$$\hat{\mathbf{S}}_{n,x} \begin{cases} \leq \hat{S}_{n,x}^{max} \times \mathbf{P}_{n,x}, & \text{if } cost_{n,x}^P \geq 0 \\ \leq \hat{S}_{n,x}^{max}, & \text{otherwise} \end{cases} \quad \forall (n, x) \in (N, X)_{store} \quad (3.15)$$

- Maximum charge/discharge

$$\mathbf{E}_{n,x,t,c}^+ \leq \hat{\mathbf{E}}_{n,x} \forall (n, x, c, t) \in (N, X, T, C)_{prod,store} \quad (3.16)$$

$$-1 \times \mathbf{E}_{n,x,t,c}^- \leq \hat{\mathbf{E}}_{n,x} \forall (n, x, c, t) \in (N, X, C, T)_{con,store} \quad (3.17)$$

Table 3.8 Valid parameters for constraining a transmission technology.

Parameter	Calliope Syntax	Thesis syntax	Default	Domain
<i>Verbose syntax</i>	<i>Calliope Syntax</i>	<i>Thesis syntax</i>		
Maximum installed energy capacity	energy_cap_max	$\hat{E}^{max}$	$\infty$	$[0, \infty]$ kW
Energy efficiency	energy_eff	$\eta^E$	1.0	$[0, \infty)$
Energy efficiency per distance	energy_eff_per_distance	$\eta^{E,pd}$	1.0	$[0, \infty)/m$
Distance	distance	distance	0	$[0, \infty]$ m
One way transmission	one_way	one_way	{0, 1}	



Figure 3.10 Calliope transmission technology.

**Transmission** Transmission technologies are those which consume an energy carrier from within a node and produce the same energy carrier within a different node (Figure 3.10, Table 3.8). Usually, constraints are duplicated for a link, to allow transmission in both directions. However, if the link is defined as *one\_way*, flow will only be allowed in the direction  $n \rightarrow n_r$  for a given link  $(n, n_r)$ . Specific constraints to apply are:

- Remote node linking

$$\hat{\mathbf{E}}_{n,x} = \hat{\mathbf{E}}_{n_r,x} \forall x \in X_{trans}, \forall (n, n_r) \in L \quad (3.18)$$

- Production/Consumption linking

$$-1 * \mathbf{E}_{n,x,t,c}^- \times (\eta_{(n,n_r),x,t}^E \times \eta_{(n,n_r),x,t}^{E,pd} \times distance_{(n,n_r),x}) = \mathbf{E}_{n_r^+,x,t,c}^+ \forall (x,t,c) \in (X,T,C)_{trans}, \forall (n,n_r) \in L \quad (3.19)$$

- Maximum transmitted energy

$$\mathbf{E}_{n,x,t,c}^+ \leq \hat{\mathbf{E}}_{n,x} \forall (n,x,t,c) \in (N,X,T,C)_{prod,trans} \quad (3.20)$$

$$-1 \times \mathbf{E}_{n,x,t,c}^- \leq \hat{\mathbf{E}}_{n,x} \forall (n,x,t,c) \in (N,X,T,C)_{con,trans} \quad (3.21)$$

Table 3.9 Valid parameters for constraining a demand technology.

Parameter	Calliope Syntax	Thesis syntax	Default	Domain
Verbose syntax				
Require resource consumption	force_resource	$R^{force}$	0	{0, 1}
Required resource	resource	R	0	$[-\infty, 0]$ kWh/m <sup>2</sup>
Maximum installed collector area	resource_area_max	$R^{area,max}$	$\infty$	$[0, \infty]$ m <sup>2</sup>
Resource scale	resource_scale	$R^{scale}$	1.0	$[0, \infty)$

**Demand** Demand ‘technologies’ are inverse supply technologies, with a smaller set of available parameters with which they can be constrained (Table 3.9). Note that the resource parameter R is negative for demand technologies. In most cases, the binary variable  $R^{force}$  is set to 1, requiring all the resource requirements of the demand technology to be met. However, instances exist where there may be a demand technology for which the demand need not be met, but perhaps there is a financial benefit to meeting some of the demand. Specific constraints to apply are:

- Required resource

$$\frac{\mathbf{E}_{n,x,t,c}^-}{\eta_{n,x,t}^E} \begin{cases} = R_{n,x,t} \times R^{scale}, & \text{if } R_{n,x}^{force} = 0 \\ \leq R_{n,x,t} \times R^{scale}, & \text{if } R_{n,x}^{force} = 1 \end{cases} \forall (n,x,t,c) \in (N,X,T,C)_{demand} \quad (3.22)$$

Where  $R^{force}$  is a binary parameter dictating whether a technology is dispatchable (e.g. diesel fuel) or not (e.g. un-curtailed PV).

- Required resource per unit area

$$\frac{\mathbf{E}_{n,x,t,c}^-}{\eta_{n,x,t}^E} \begin{cases} = R_{n,x,t} \times R^{scale} \times \hat{\mathbf{R}}_{n,x}^{area}, & \text{if } R_{n,x}^{force} = 0 \\ \leq R_{n,x,t} \times R^{scale} \times \hat{\mathbf{R}}_{n,x}^{area}, & \text{if } R_{n,x}^{force} = 0 \end{cases} \quad (3.23)$$

$$\forall (n, x, t, c) \in (N, X, T, C)_{demand, area}$$

**Node** Two constraints act across an entire node:

- System balance

$$\sum_{x \in X} (\mathbf{E}_{n,x,t,c}^+ + \mathbf{E}_{n,x,c,t}^- + \mathbf{E}_{n,x,c,t}^{ex}) + \mathbf{slack}_{n,c,t}^+ + \mathbf{slack}_{n,c,t}^- = 0 \forall (n, t, c) \in (N, T, C) \quad (3.24)$$

- Available area

$$\sum_{x \in X_{area}} \mathbf{R}_{n,x}^{area} \leq available\_area_n \forall n \in (N)_{area} \quad (3.25)$$

## Objective function

Table 3.10 Valid parameters for applying costs to decision variables in Calliope.

Parameter	Calliope Syntax	Thesis syntax	Domain
<i>Verbose syntax</i>			
Cost of energy capacity	energy_cap	cost <sup>Ê</sup>	(-inf, inf)/kW
Cost of energy capacity, per unit distance	energy_cap_per_distance	cost <sup>Ê,pd</sup>	(-inf, inf)/kW/m
Purchase cost	purchase	cost <sup>P</sup>	(-inf, inf)
Cost of resource collector area	resource_area	cost <sup>Ê<sup>area</sup></sup>	(-inf, inf)/m <sup>2</sup>
Cost of storage capacity	storage_cap	cost <sup>Ŝ</sup>	(-inf, inf)/kWh
Carrier export cost	export	cost <sup>E<sup>ex</sup></sup>	(-inf, inf)/kWh
Carrier consumption cost	om_con	cost <sup>E<sup>-</sup></sup>	(-inf, inf)/kWh
Carrier production cost	om_prod	cost <sup>E<sup>+</sup></sup>	(-inf, inf)/kWh
Interest rate	interest_rate	rate <sup>int</sup>	(-inf, inf)
Technology lifetime	lifetime	L	[0, ∞) years

Costs given in Table 3.10 can be applied to the aforementioned decision variables, for use in the objective function. The objective function is the sum of annualised investment costs and operational costs for a given cost class  $k$  (such as ‘monetary’ or ‘carbon’):

$$\begin{aligned}
minTSCF \times \sum_{n,x} & \left( \hat{\mathbf{E}}_{n,x} \times cost_{n,x,k}^{\hat{\mathbf{E}}} + \right. \\
& \hat{\mathbf{E}}_{n,x \in X_{trans}} \times cost_{n,x,k}^{\hat{\mathbf{E}},pd} \times distance_{(n,n_r),x} + \\
& \hat{\mathbf{R}}_{n,x}^{area} \times cost_{n,x,k}^{\hat{\mathbf{R}}^{area}} + \\
& \left. \hat{\mathbf{S}}_{n,x} \times cost_{n,x,k}^{\hat{\mathbf{S}}} + \mathbf{P}_{n,x} \times cost_{n,x,k}^{\mathbf{P}} \right) \\
& + \sum_{n,x,c,t} \left( \mathbf{E}_{n,x,c,k,t}^+ \times cost_{n,x,k,t}^+ - \mathbf{E}_{n,x,c,k,t}^- \times cost_{n,x,k,t}^- + \mathbf{E}_{n,x,c,k,t}^{ex} \times cost_{n,x,k,t}^{ex} \right) \\
& + bigM \sum_{n,c,t} (\mathbf{slack}_{n,c,t}^+ - \mathbf{slack}_{n,c,t}^-)
\end{aligned} \tag{3.26}$$

Where  $TSCF$  is the timeseries scaling factor  $\frac{T}{8760} \times \frac{rate^{int}}{1-(1+rate^{int})^{-L}}$ , to scale the investment costs to the model time horizon  $T$ , and  $bigM$  is a value sufficiently large such that slack variables are only utilised to ensure model feasibility (e.g.  $1 \times 10^{10}$ ). Any value of unmet demand or unused supply implies a model that has not been correctly formulated. Without slack variables, the linear optimiser would return the model as infeasible with no accompanying indication as to why, making it difficult for a modeller to address the issue.

### 3.2.2 Data entry and internal workflow

The user interface for Calliope is designed to be easy to understand. As such, the ‘‘human friendly data serialization standard for all programming languages’’, YAML, is used for model formulation (YAML, 2009). Timeseries data is stored in a comma separated value format and referred to where necessary in the model formulation. The code block in Figure 3.11 shows how a CHP technology would be defined in YAML, under the ‘techs’ heading. Other headings under which the model is formulated are ‘locations’ (for allocating technologies to nodes), ‘links’ (for linking nodes), ‘model’ (for model-wide settings), and ‘run’ (for optimisation solver settings).

Once the model has been prepared, state-of-the-art Python3 libraries are used to process the data for sending to the solver. Figure 3.12 shows the internal workflow of Calliope, in which ruamel.yaml (Neut, 2018), pandas (McKinney, 2010), xarray (Hoyer and Hamman, 2017), and pyomo (Hart *et al.*, 2017) are key to preparing the model. An xarray Dataset is the format in which the final processed data is stored. It can handle multi-dimensional arrays and array-specific attributes. The Dataset can be saved in the Network Common Data Format (NetCDF), which is a file format specific to sharing scientific data. The solver software itself can be chosen by the modeller, from the full set made available by pyomo. This allows the same model to be solved seamlessly with different software. Indeed, tests are in place to ensure that solver results from IBM ILOG CPLEX, Gurobi, GPLK, and CBC are consistent

```

techs:
  chp:
    essentials:
      name: "Combined heat and power"
      color: "#E4AB97"
      # `conversion_plus` technologies allow multi-energy conversion
      parent: conversion_plus
      # The primary carrier out is used to apply constraints and costs
      # such as energy capacity and carrier production costs
      primary_carrier_out: electricity
      carrier_in: gas
      carrier_out: electricity
      # heat is given in a separate stream to electricity,
      # so heat and electricity output are linked
      carrier_out_2: heat
    constraints:
      export_carrier: electricity
      energy_cap_max: 1500
      energy_eff: 0.405
      # heat to power ratio = 0.8
      carrier_ratios.carrier_out_2.heat: 0.8
      lifetime: 25
    costs:
      monetary:
        interest_rate: 0.10
        energy_cap: 750
        # .4p/kWh for 4500 operating hours/year
        om_prod: 0.004
        # timeseries export price found in CSV file
        export: file=export_power.csv
        # A carbon cost for consuming gas is included
      carbon:
        om_con: 0.196

```

Figure 3.11 Example of Calliope YAML model formulation in which a **CHP** technology is defined.

for two example Calliope models. Commercial solvers offer the best solution times, as seen previously in Table 3.1, but access to such solvers is often limited to the academic community, where their use is free.

### 3.2.3 Timeseries data and clustering

Timeseries data is loaded into the model in discrete time steps, stored in comma-separated value (CSV) format. Many parameters can vary in time, such as energy demand and intermittent energy supply. Additionally, technology efficiencies and costs/revenues may be given as static or time-varying values. There is no time horizon nor resolution requirement on the data, provided all datasets are consistent. A timestamp is associated with a value for a parameter, with the resolution of that timestamp given by the time difference between it and the following timestamp in the series. The generalised timeseries input structure allows a user to:

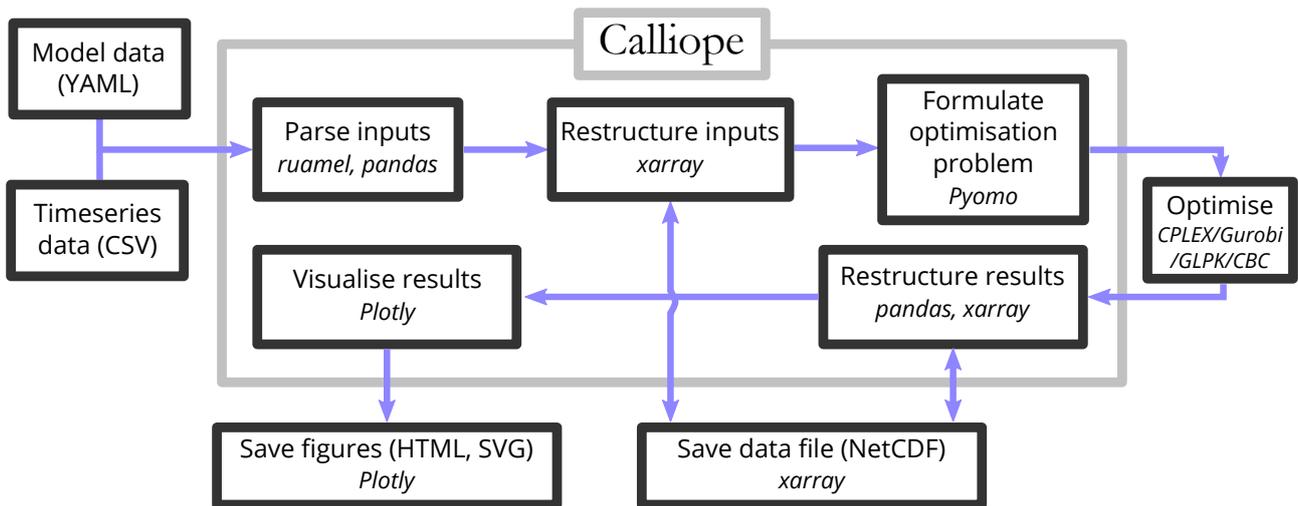


Figure 3.12 Internal Calliope workflow, including third party library dependencies.

- define data at any resolution, be it every second, hour, or year;
- provide more data than is necessary, then select a subset of the timeseries over which to optimise the model;
- include different timestep resolutions within the same timeseries, e.g. three hour resolution overnight and one hour resolution in the day, and;
- load a highly resolved timeseries, then use Calliope functions to reduce the size of the time dimension for tractable optimisation.

The size of the time dimension can be reduced by downsampling and clustering. As discussed in Section 2.1.1, downsampling aggregates timesteps whilst clustering chooses **typical days (TDs)** to represent the full timeseries. In either case, outlying days can be masked prior to timeseries reduction, to ensure information on those days is kept intact. If a user has access to highly resolved data, algorithmic clustering circumvents the laborious task of selecting the typical days that lead to the least error in the model. Calliope can algorithmically cluster using agglomerative hierarchical and k-means methods. In both cases, all timeseries data is normalised and all data associated with a given date is concatenated to create daily ‘observations’. Each observation has a length equal to the number of timesteps in a day multiplied by the number of parameters defined as a timeseries. This allows similar days to be chosen based on the similarity across all timeseries data. Alternatively, a user can select a subset of timeseries parameters with which to select **TDs**. For instance, clustering could take place using only weather data, ignoring energy demand data.

Agglomerative hierarchical clustering begins with every observation in its own cluster. Clusters move up the hierarchy by forming pairs which minimise a metric, such as the Euclidean distance. Clustering terminates when either a sufficiently small number of clusters have been chosen or the distance between forming pairs begins to increase above a given threshold, either of which is predefined. This approach was formalised for use on timeseries data by Nahmmacher *et al.* (2016) and shown to be as effective as k-means clustering for certain energy system model configurations (Kotzur *et al.*, 2018a; Pfenninger, 2017a).

K-means clustering randomly assigns  $k$  cluster centroids to the data, where  $k$  is the predefined number of desired clusters. Each centroid represents as much data as it is closer to than surrounding centroids. The positions of centroids are then moved sequentially by a heuristic algorithm, to further minimise the sum of the distance of the data they represent (the inertia). The centroids will converge on representing clusters where the Euclidean distance of all observations from their respective centroids is minimised. But they may become locked in local minima, causing incorrect clustering. This effect can be mitigated by running the algorithm several times for the same dataset, each time with a new random allocation of initial centroids. Clustering a full year of energy system timeseries data to five TDs 1000 times, Pfenninger (2017a) found that k-means was sufficiently stable to be used without concern regarding local minima.

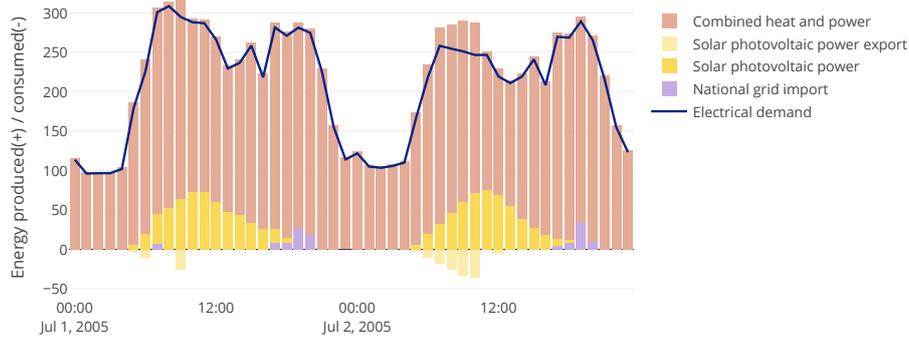
Unlike in hierarchical clustering, the process of converging on clusters cannot be tracked with k-means. Indeed, the choice of  $k$ , the number of resulting TDs following clustering, may not satisfactorily encompass all underlying clusters in the data (assuming that there are any). Rather than select  $k$  subjectively, it is possible to quantify the ‘quality’ of resulting clusters from iteratively greater  $k$ .  $k$  is found once the quality of clusters reaches a certain threshold. Many thresholds exist, the best of which is ‘Hartigan’s’ rule (Chiang and Mirkin, 2010). Hartigan and Hartigan (1975) proposed that the best choice of  $k$  can be found by calculating the difference in total inertia when increasing  $k$  to  $k + 1$ :  $HK = (\frac{W_k}{W_{k+1}} - 1)(N - k - 1)$ , where  $N$  is the total number of observations and  $W$  is the total inertia. Large decreases in inertia between  $k$  and  $k + 1$  suggest that the ‘right’ number of clusters has not yet been converged upon. Once  $HK$  decreases below 10, the lowest  $k$  is considered to be the right number of clusters.

### 3.2.4 Visualising results

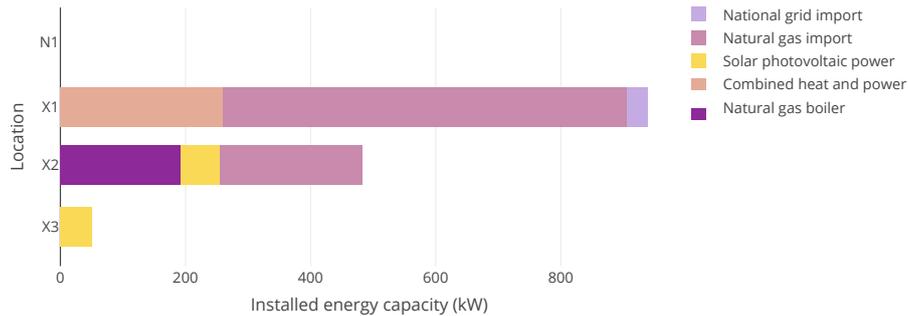
The interactive plotting library Plotly (Plotly, 2015) is used to visualise results such as the operation schedule, optimal network routing, and technology capacities (Figure 3.13). Unlike most plotting libraries used by similar modelling frameworks, the interactivity of the figures allows a user to select subsets of data to hide, use a dropdown menu to switch between results, and zoom in on areas of interest in a particularly complex plot. Thus, the dissemination of data is straightforward, for both seasoned modellers and casual users.

### 3.2.5 Ongoing development

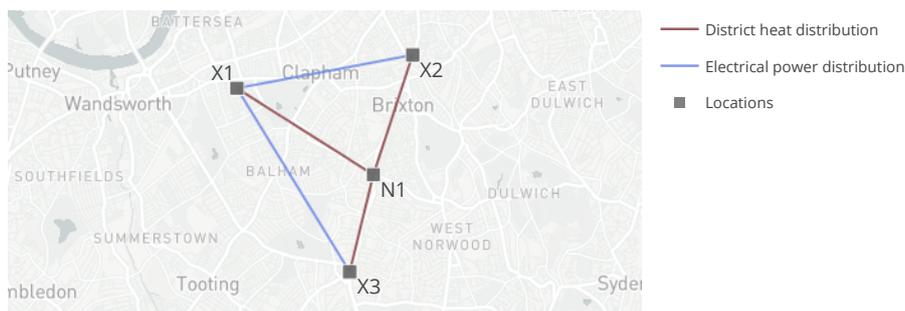
As any open-source project, Calliope continues to grow. Equally, software improvements are necessarily slow, to ensure they are well tested, well documented, and usable by a wide audience. As such, there is an extensive roadmap for Calliope. Indeed, functionality is introduced in each of the following chapters, beyond that described in this chapter. Additional functionality is still accessible, via branches of the main repository, but is not yet officially packaged in the primary release of the software. Ongoing development includes:



(a) Electricity carrier operation schedule.



(b) Technology energy capacity.



(c) Transmission network.

Figure 3.13 Figures produced using Calliope plotting functions. Plots are interactive when viewed as HTML.

- piecewise linearisation of part-load curves and technology costs, using both [special order set constraint of type 2 \(SOS2\)](#) and [convex bounding sets \(CBS\)](#) (Chapter 4);
- risk-neutral and risk averse scenario optimisation (Chapters 5 and 6);
- objective functions which allow multiple cost classes to be included at once (Chapter 6), and;
- [out-of-sample](#) testing with differences in the data available in the timestep window compared to the timestep horizon (Chapter 5).

### 3.3 Open source district energy optimisation

Although energy system models are generally moving towards openness, as discussed in Section 2.2.3, those that focus on the district scale have been lagging behind. Calliope began as an open-source endeavour and its development to correctly capture the intricacies of district scale models has arguably realised a more robust open-source model across all scales. To assess this improvement, six recommendations for successful open-source energy models are considered (DeCarolis *et al.*, 2012):

1. Make source code publicly accessible.
2. Make model data publicly accessible.
3. Make transparency a design goal.
4. Utilize free software tools.
5. Develop test systems for verification exercises.
6. Work toward interoperability among models.

Calliope meets all six of these recommendations and four have been improved in parallel with the introduction of district scale operability. Particular improvements which have been realised are:

1. *Data accessibility* A greater number of data streams have been necessarily accessed to develop district-scale models. Once studied, they have been made publicly available, increasing the model data accessible for direct use with the Calliope framework.
2. *Model transparency* In both the user interface and internal workflow, transparency of model code has been reconfigured to be a central focus. Indeed, when the development team for a piece of software grows, transparency must inevitably improve. The internal workflow is now more accessible to an individual external to the project, ensuring broader scrutiny of model practices.
3. *Testing* The source code test coverage has increased from less than 70% to nearly 95%. In so doing, Calliope is demonstrably the most well-tested open-source energy system model available. Tests cover individual functions within the source code as well as tracking a number of illustrative models from YAML entry to HTML visualisation. The optimisation process itself can become incredibly opaque, so verifying that each constraint is applying the expected bounds on decision variables is vital to ensuring model outputs are trustworthy.
4. *Interoperability* Model interoperability has improved as a direct consequence of the refactorisation of Calliope for improved transparency. The NetCDF data structure that describes a model can be easily shared between modellers and used to connect directly

to other models. For instance, demand simulation models can inform the demand 'resource' values, while outputs can be directly linked to optimal control models. Similar to the climate science community, Calliope is championing the move towards NetCDF being a standard data format for energy system models. If achieved, the myriad of existing energy models will share a common data structure for the cross-validation of models.

# Chapter 4

## Exploring the linearisation of ‘reality’

Work in this chapter pertaining to comparisons with a metaheuristic algorithm were undertaken in collaboration with Shintaro Ikeda (University of Tokyo). Shintaro developed the  $\epsilon$ DE algorithm and collated the nonlinear technology part-load data. Parts of this chapter have been published as:

B. Pickering *et al.* (2016). ‘Comparison of Metaheuristic and Linear Programming Models for the Purpose of Optimising Building Energy Supply Operation Schedule’. In: *12th REHVA World Congress*. Vol. 6. Aalborg, Denmark: Aalborg University, Department of Civil Engineering. ISBN: 87-91606-31-4. URL: [http://vbn.aau.dk/files/233775414/paper\\_529.pdf](http://vbn.aau.dk/files/233775414/paper_529.pdf) (visited on 11/08/2016)

B. Pickering and R. Choudhary (July 2017). ‘Applying Piecewise Linear Characteristic Curves in District Energy Optimisation’. In: *Proceedings of the 30th International Conference on Efficiency, Cost, Optimisation, Simulation and Environmental Impact of Energy Systems*. San Diego, USA, pp. 1080–1092

Technology characteristics are inherently nonlinear, some more so than others. When modelling an energy system as a set of linear constraints, the reality of technology operation is therefore lost. Metaheuristic algorithms are well placed to handle nonlinearities in a model. However, there are limitations in their ability to represent large district energy systems, particularly if getting close to a global optimum is desirable. Indeed, most district energy system studies are chosen to fit within a linearised representation of reality (Sameti and Haghghat, 2017). Two prominent areas in which linearisation is undertaken are 1) technology operating variability due to temperature fluctuations and 2) operating efficiency at part-load. Nonlinear part-load operation has been observed in empirical studies (Best *et al.*, 2015) and can be the recommended representation of technologies given by a government body (Ikeda and Ooka, 2016). In some cases, the expected deviation between the ‘realistic’<sup>1</sup> and linearised representation of a technology leads to certain characteristics or entire technologies being removed from consideration in a study (Jennings *et al.*, 2014; Steen *et al.*, 2015).

Aware of nonlinearities, but not wanting to ignore them, some studies have begun considering piecewise linearisation (Bischi *et al.*, 2014; Evins *et al.*, 2014; Wakui and Yokoyama, 2014). The nonlinear partial load curve of technologies can be separated into linear sections,

---

<sup>1</sup>No computational representation of technology operation will truly capture its realistic characteristics. In this thesis ‘realistic’ is referred to as the best-case representation of technologies.

with each section being linear. The greater the number of sections, the closer the representation of nonlinearity. Bischi *et al.* (2014) showed that piecewise linearisation increases computation time, but there is no change in result if using more than 10 pieces to describe a curve. However, Bischi *et al.* (2014) did not examine how well piecewise curves converge on the 'real' nonlinear representation. In the power electronics sector, del Valle *et al.* (2009) compared enhanced particle swarm optimisation (a metaheuristic method) to **mixed integer linear programming (MILP)** (referred to as the branch & bound method). The results indicated that particle swarm optimisation was better than branch & bound optimisation in both time to solution and **objective function value (OFV)**. However, the particle swarm optimisation inputs were linearised for comparability with the MILP model. Indeed, direct comparison between nonlinear models and their linearised counterparts is not generally undertaken, certainly not in energy system modelling.

In this chapter the impact of linearisation on energy systems is tested. To do so, nonlinear characteristic curves of various technologies are taken from Japanese government guidelines (The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE), 2003). The curves describe the energy consumption of technologies at part-load rates, i.e. when a technology is not operating at full capacity. A metaheuristic model was used to provide a nonlinear benchmark for an energy system, against which various linear models were tested. In Section 4.1 piecewise linearisation methods are described before applying them to two case studies, one in Japan and the other in the UK (both detailed in Section 4.2). The impact of linearisation on objective function value, actual operation cost (as simulated ex-post), and investment/dispatch decisions is given in Section 4.3.1. Exploring piecewise linearisation further, Section 4.3.2 looks at how a pre-optimisation step could be used to reduce computational effort in the subsequent MILP models without loss of nonlinear curve fitting.

## 4.1 Linearising part-load curves

When using a nominal value for a technology characteristic, the constraint generated with respect to that value is of the form  $y = mx$ . For instance, a **coefficient of performance (COP)** of 4 for a heat pump dictates that its heat output ( $y$ ) is 4 ( $m$ ) times the electricity input ( $x$ ). The linear relationship between the decision variables  $y$  and  $x$  allows us to undertake linear programming. When a technology characteristic is to be described by anything more complex than a nominal value, we enter into nonlinear constraints. Although linear when  $x$  assumes a value greater than zero, Figure 4.1a displays a constraint with a clear discontinuity at  $x = 0$ , where  $y$  is simultaneously 0 and  $C$ . In linear programming, binary decision variables can be used to overcome this discontinuity. The binary constraint creates a switch for the constraint, such that  $y = 0$  if  $x = 0$  and  $y = mx + C$  otherwise. The curvature of the remaining characteristic curves in Figure 4.1 precludes the use of a single binary constraint. When faced with characteristic curves which are not straight lines in MILP, they must be described by a set of straight lines connected by breakpoints, i.e. piecewise linearised (Figure 4.1c). Breakpoints are positioned at the intersection of two straight lines along piecewise

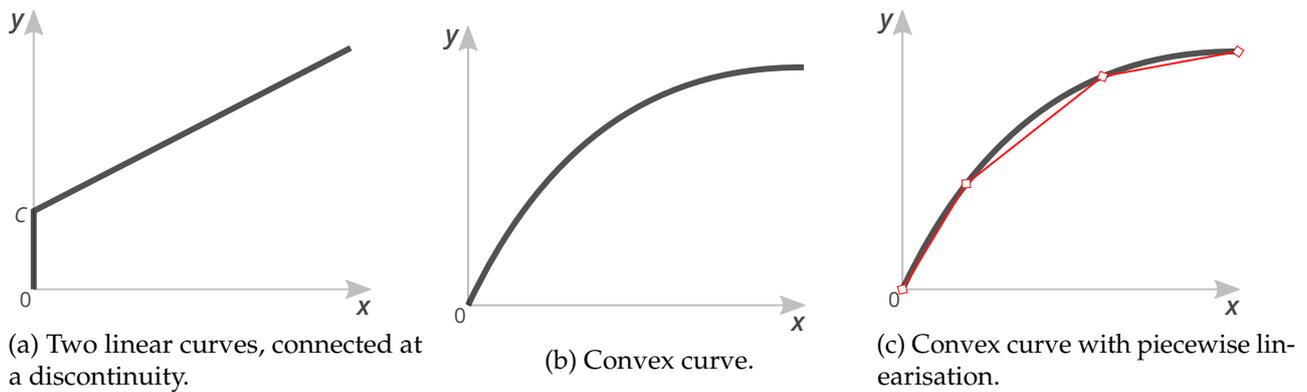


Figure 4.1 Example of nonlinear curves, including a piecewise representation.

curves. They are positioned directly on the nonlinear curve, thus accurately represent the nonlinear curve at that point. Between breakpoints, the straight lines will, in most cases, deviate from the nonlinear curve. Two approaches can be taken in piecewise linearisation: [convex bounding sets \(CBS\)](#) and [special order set constraint of type 2 \(SOS2\)](#). Models with piecewise representation of a technology use the [SOS2](#) methodology (Bischi *et al.*, 2014; Milan *et al.*, 2015; Wakui *et al.*, 2014). However, the number of pieces used varies between models. Bischi *et al.* (2014) found that up to ten pieces shows minimal increased computational time compared to the complete linear case, which is sufficient to describe most of the nonlinear curves encountered in energy systems.

### Convex bounding sets

It would also be valid to impose the constraint  $y \leq mx$  instead of  $y = mx$  on the aforementioned example of a heat pump COP as it is suboptimal for  $y$  to sit anywhere other than on the line described by  $mx$ . If producing less heat, there is no cost-optimal reason to continue consuming more electricity than given by  $y/m$ . Yet, the solution space for variable  $y$  has now expanded from being on the line  $y = mx$  to include everything below it, i.e. the shaded region in Figure 4.2a. If a second linear constraint is added, now described by  $y \leq mx + C$ , the shaded region available to  $y$  can be changed, still with the knowledge that it will not sit anywhere *but* on the line now described by two curves (Figure 4.2b). This premise of defining a convex hull solution space by several constraints is how a linear optimisation problem is constructed, albeit in multidimensional space, as already detailed in Section 3.1. By continuing to add constraints of the form  $y \leq mx + C$ , the description of any convex characteristic curve can be converged upon. If the example heat pump has a heat output to power input relationship as given in Figure 4.2c, CBS can be used to estimate the curve. Of course, only with an infinitely large set of constraints can a smooth convex curve truly be described. Additionally, there are two strict requirements for this method to function, both related to the inequality used in each of the set of constraints. First, the nonlinear constraint must have a strictly increasing/decreasing gradient. Figure 4.2d shows what will happen if this is not the case. Certain lines describing the piecewise curve encroach on the search space where they should not. The result is an upper bound of the search space that does not resemble the original nonlinear curve. Second, it must be suboptimal for the decision

variable  $y$  to be anywhere other than on the line described by the upper bound of the convex hull. If  $y$  was instead the cost of consuming electricity, then it would actually be optimal for it to be always zero, which is within the search space. Thus, decision variables which are optimal at their upper bound can only be described by a curve that has a strictly decreasing gradient. Conversely, decision variables which are optimal at their lower bound can only be described by a curve that has a strictly increasing gradient. **CBS** are clearly limited in use, but they allow a model to remain purely linear, or limit the constraint to requiring only one binary constraint (if the nonlinear curve does not reach the origin). This method can also be extended easily to three dimensional surfaces, where the constraints are of the form  $f(x, y) = my + Cx$ , with  $f(x, y)$ ,  $y$  and  $x$  all being decision variables.

### Special order set constraint of type 2

Where convex bounding sets fail to describe a particular nonlinear curve, it is possible to use **SOS2** instead. First introduced by Beale and Tomlin (1970), **SOS2** are "a set of consecutive variables in which not more than two adjacent members may be non-zero in a feasible solution". Consider a curve similar to that given in Figure 4.2d, now described as the function  $f(x)$  in Figure 4.3a. Points  $P_1, \dots, P_K$  are assigned along the curve, with coordinates  $(\hat{x}_k, f(\hat{x}_k))$ ,  $k = 1, \dots, K$ . Any point  $x$  in the closed interval  $[\hat{x}_k, \hat{x}_{k+1}]$  may be written as

$$x = \alpha_k \hat{x}_k + \alpha_{k+1} \hat{x}_{k+1} \quad \text{where} \quad \alpha_k + \alpha_{k+1} = 1 \text{ and } \hat{x}_k, \hat{x}_{k+1} \geq 0. \quad (4.1)$$

As  $f(x)$  is linear in the interval, it can be written as

$$f(x) = \alpha_k f(\hat{x}_k) + \alpha_{k+1} f(\hat{x}_{k+1}). \quad (4.2)$$

which leads to the representation of  $f(x)$  using a set of weighting variables (known as the *special ordered set type two variables*),  $\alpha_k$ ,  $k = 1, \dots, K$ , by the equality

$$f(x) = \alpha_1 f(\hat{x}_1) + \alpha_2 f(\hat{x}_2) + \dots + \alpha_K f(\hat{x}_K) \quad (4.3)$$

where  $\alpha_1 \hat{x}_1 + \alpha_2 \hat{x}_2 + \dots + \alpha_K \hat{x}_K - x = 0$ ,  $\alpha_k \geq 0$  and  $\alpha_1 + \alpha_2 + \dots + \alpha_K = 1$ ,  $k = 1, \dots, K$ . With the additional condition that not more than two adjacent variables can be non-zero any one time, it is possible to represent each point on the piecewise curve given in 4.3a within a **MILP** model. This does mean that we require a reasonable number of binary variables for every characteristic curve, making the representation more complex than with **CBS**, which can be described completely in a **linear programming (LP)** model. Thus, solution times can be expected to be greater when using **SOS2**, even up to the point of causing intractability. As with **CBS**, nonlinear curves linking more than two decision variables can be piecewise linearised using special ordered sets. The most common approach is to have  $n$  breakpoints  $x_1, \dots, x_n$  on the  $x$  axis and  $m$  sampling points  $y_1, \dots, y_m$  on the  $y$  axis (D'Ambrosio *et al.*, 2010).  $f(x, y)$  is evaluated for each breakpoint. Any point  $(\hat{x}, \hat{y})$  can be evaluated within the rectangle bounded by  $(x_i, y_j)$ ,  $(x_{i+1}, y_j)$ ,  $(x_i, y_{j+1})$ , and  $(x_{i+1}, y_{j+1})$ , which contains two triangles created by its diagonal  $[(x_i, y_j), (x_{i+1}, y_{j+1})]$  (Figure 4.3b). By convex combination

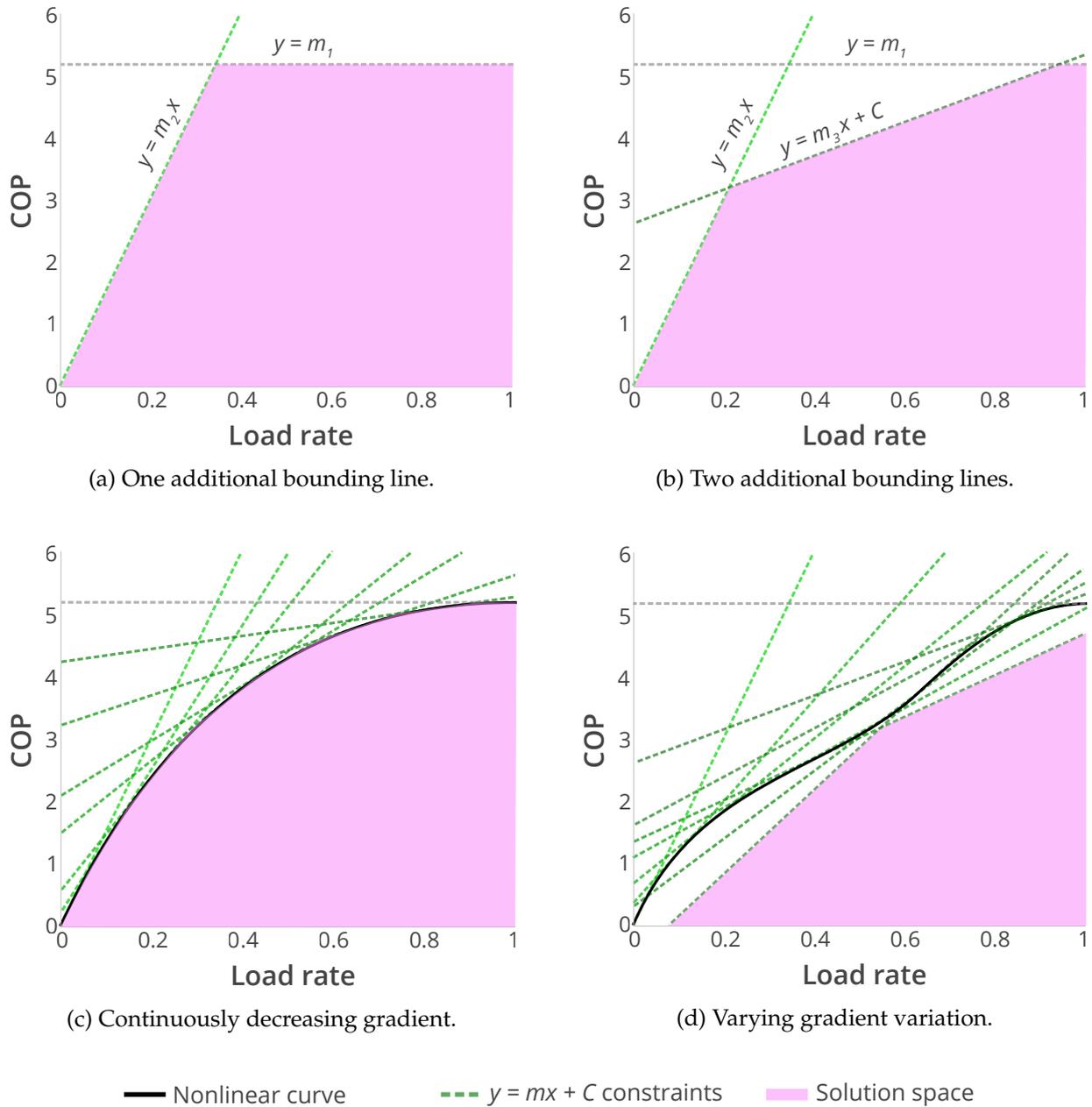
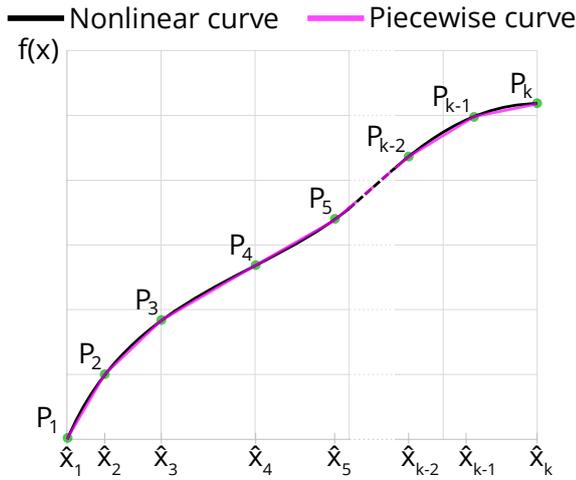
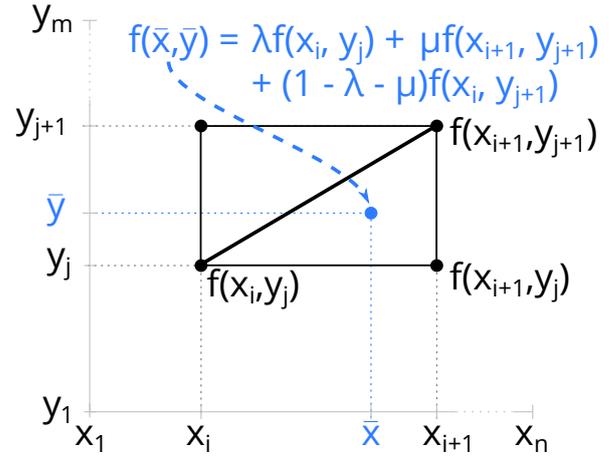


Figure 4.2 Application of CBS to describe a nonlinear curve. (a) shows a single bounding line, extended to two bounding lines in (b). In (c) the bounding set is large enough to describe the nonlinear curve well. Due to the gradient not strictly increasing/decreasing, (d) shows ineffectiveness of CBS when applied to a more complex curve.

of the function, values evaluated at the vertices of the triangle containing  $(\hat{x}, \hat{y})$ ,  $f(\hat{x}, \hat{y})$  can be ascertained.



(a) Graphical representation of SOS2 piecewise linearisation.  $f(x)$  is the sum of weighted values  $\alpha_i f(x_i)$  and  $\alpha_{i+1} f(x_{i+1})$ , with all other values of  $\alpha$  being zero.



(b) Graphical representation of 3D piecewise linearisation.  $f(\hat{x}, \hat{y})$  is the sum of weighted decision variables  $\lambda$  and  $\mu$  applied to  $f(x_i, y_j)$ ,  $f(x_{i+1}, y_{j+1})$  and  $f(x_{i+1}, y_j)$ .

Figure 4.3 Graphical representation of piecewise linearisation using SOS2.

### Measuring Impact

**Ex-ante linearisation error** Between breakpoints on a piecewise curve, the straight line will usually deviate from the nonlinear curve. By measuring this deviation along the length of the nonlinear curve, we can assess the linearisation error of the curves ex-ante. The goodness-of-fit of two curves can be compared using **root-mean-square error (RMSE)**, which discretises the curves (in this case into 1000 points) and measures the error at each point using the equation:

$$RMSE = \sqrt{\left( \frac{\sum_{p \in P} (y_{1p}^2 - y_{2p}^2)}{length(p)} \right)} \tag{4.4}$$

where points  $p$  are the discretised curve points in the set  $P$ ,  $y_1$  and  $y_2$  are the curve  $y$ -axis values for the nonlinear and linear curve, respectively. Ex-ante linearisation error can be assessed for single line representations of the curve as well as piecewise curves. In the case of piecewise curves, a set of breakpoints will have at least one set of positions which minimises the ex-ante linearisation error. Although we could make a good guess at optimal placement, the process of locating breakpoint sets optimally can be simplified by automation. Setting the objective as the piecewise curve that has the least error relative to the nonlinear curve allows us to conduct computational optimisation in this preprocessing step. Additionally, the constraint that the gradient must be strictly increasing or decreasing can be applied, creating piecewise curves which meet the requirements set out in section 4.1.

Breakpoint allocation can be undertaken during model pre-processing, by parameter optimisation. Previous studies have used heuristic algorithms to piecewise linearise (Horner

and Beauchamp, 1996; Siriruk, 2012). In this study, the heuristic [sequential least squares programming \(SLSQP\)](#) method (Kraft *et al.*, 1988) has been used with the objective function to minimise the [RMSE](#) between each nonlinear curve and its piecewise counterpart. As heuristic methods do not necessarily reach a global optimum, the [SLSQP](#) optimisation must be run multiple times to improve the chances of a useful result. In each run, the breakpoint positions are randomly placed on the curve at the start of the optimisation.

**Ex-post linearisation error** An optimised energy system model will provide a decision maker with a recommended investment portfolio and operation schedule. If the model is based on linearised inputs, there will be a difference in ‘real’ operation compared to that given by the model. Depending on the shape of the curve being linearised, the difference may be negative or positive. A positive impact is caused by a *pessimistic* linearisation: the piecewise curve leads to an average of greater consumption than its nonlinear counterpart. Conversely, a negative impact is caused by an *optimistic* linearisation. It is possible to undertake ex-post analysis on the optimal solution to measure the impact of linearisation, be it positive or negative. To do so, decision variables (technology capacity, hourly dispatch, etc.) are fixed to those which produce the minimum [OFV](#). Then, a simulation is run with the nonlinear technology characteristics in place. The linearisation error is thus the additional system cost when running the simulation when compared to the [OFV](#).

Generally, when considering part-load consumption curves, the energy balance during the simulation is undisturbed; only the magnitude of consumed resource (electricity, gas, etc.) will change. However, in a multi-energy system there are often intermediate steps, where the output of one technology acts as the input to another. A [heat recovery absorption refrigerator \(HRAR\)](#) is an example of this, as it can consume waste heat from a [combined heat and power plant \(CHP\)](#) to produce cooling. In such cases, a discrepancy in the magnitude of heat required by the [HRAR](#) will impact the quantity of heat the [CHP](#) must produce. Similarly, if the [CHP](#) has a part-load [heat-to-power ratio \(HTP\)](#) then at any given output the amount of heat it produces per unit of electricity will vary. If the linearisation is optimistic, additional energy will need to be sought from somewhere during the simulation. The most straightforward way of meeting the energy deficit is to consider a ‘slack’ technology. For example, in both the aforementioned cases an auxiliary boiler could be assumed to be available to produce heat from gas if required. If there is too much energy produced (i.e. the linearisation was pessimistic) then we assume it is dissipated to the atmosphere by some zero-cost method.

## 4.2 Model configuration

### 4.2.1 Technologies

The two case studies considered in this chapter share the same set of technologies, detailed in Table 4.1. Many of these technologies have been chosen on the existence of nonlinear curves by which to describe their part-load characteristics, as recommended for use by The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE) (2003). Shown

in Figure 4.4, part-load curves are available for an **air source heat pump (AHP)**, **electric chiller (ECh)**, **HRAR**, **absorption refrigerator (AR)**, **CHP** (including **HTP**), and **natural gas boiler (NGB)**. Pumps for all technologies that handle fluid flow also operate nonlinearly, an example of which is given by the **cold water StoreT (StoreT-C)** pump in Figure 4.4f. Many of the curves exhibit the noticeable trait of a discontinuity at low load rate which matches the ‘minimum’ load rate of the technologies, under which manufacturers recommend that the technologies do not operate. Below minimum load rate, inefficiencies in the system that are not load-dependent are more pronounced in their impact on energy consumption. Some technologies, such as the **NGB**, **AR**, and **HRAR** have an almost linear relationship above minimum load rate. Conversely, Figure 4.4a shows that the **ECh** acts nonlinearly across its entire operating range. For the **CHP**, there is a reasonable disparity between the realistic operation and rated efficiency, particularly when considering the heat to power ratio (Figure 4.4h).

Additional technologies are included beyond those with nonlinear part-load curves, to test the complexities of multi-energy systems. **Solar photovoltaic panels (PVs)** and **solar thermal panels (STs)** allow intermittent, renewable solar energy to be consumed. They do also operate nonlinearly in their conversion of solar energy into electricity and heat, respectively. However, this nonlinearity has been pre-computed such that the ‘available’ energy from either technology in each time interval is the electricity or heat output. Pre-computation of **PV** available energy was undertaken by Pfenninger and Staffell (2016) and acquired for this study from <https://renewables.ninja> (Pfenninger and Staffell, 2016). Using MERRA reanalysis data, the heat output equation proposed by Brunold *et al.* (1994) was used to pre-compute the **ST** heat output.

**electrical battery storage (StoreE)** adds an electrical temporal buffer to mirror the **thermal energy storage (StoreT)**. None of the storage technologies are considered to operate nonlinearly. Maximum flow rates for charge and discharge of the **StoreT** technologies have been limited to ensure that mixing effects are avoided. Two technologies are unique to their respective case studies: the **AR** is used only in case study 1, while the **ST** is only used in case study 2. Finally, not all information pertaining to each technology in Table 4.1 is used in every model run. Explicit exclusions are highlighted when detailing the run configuration.

## 4.2.2 Case studies

### Case study 1: Japanese hotel

This study investigated meeting the energy load of a representative Japanese hotel. The hotel has 20,000m<sup>2</sup> total floor space, with demand for electricity, hot water, and space cooling (Figure 4.5) over a 24-hour summer day. As there is only one building considered in this study, there are no distribution networks. However, this allows for greater detail in the internal energy distribution, as given in Figure 4.6. Figure 4.5 shows gas and electricity prices. Electricity export, available for the **PV** and **StoreE** technologies, produces a revenue for the system. The electricity sale price is static throughout the day, while the purchase price varies by dynamic pricing. The variable pricing is inferred from the building electricity demand

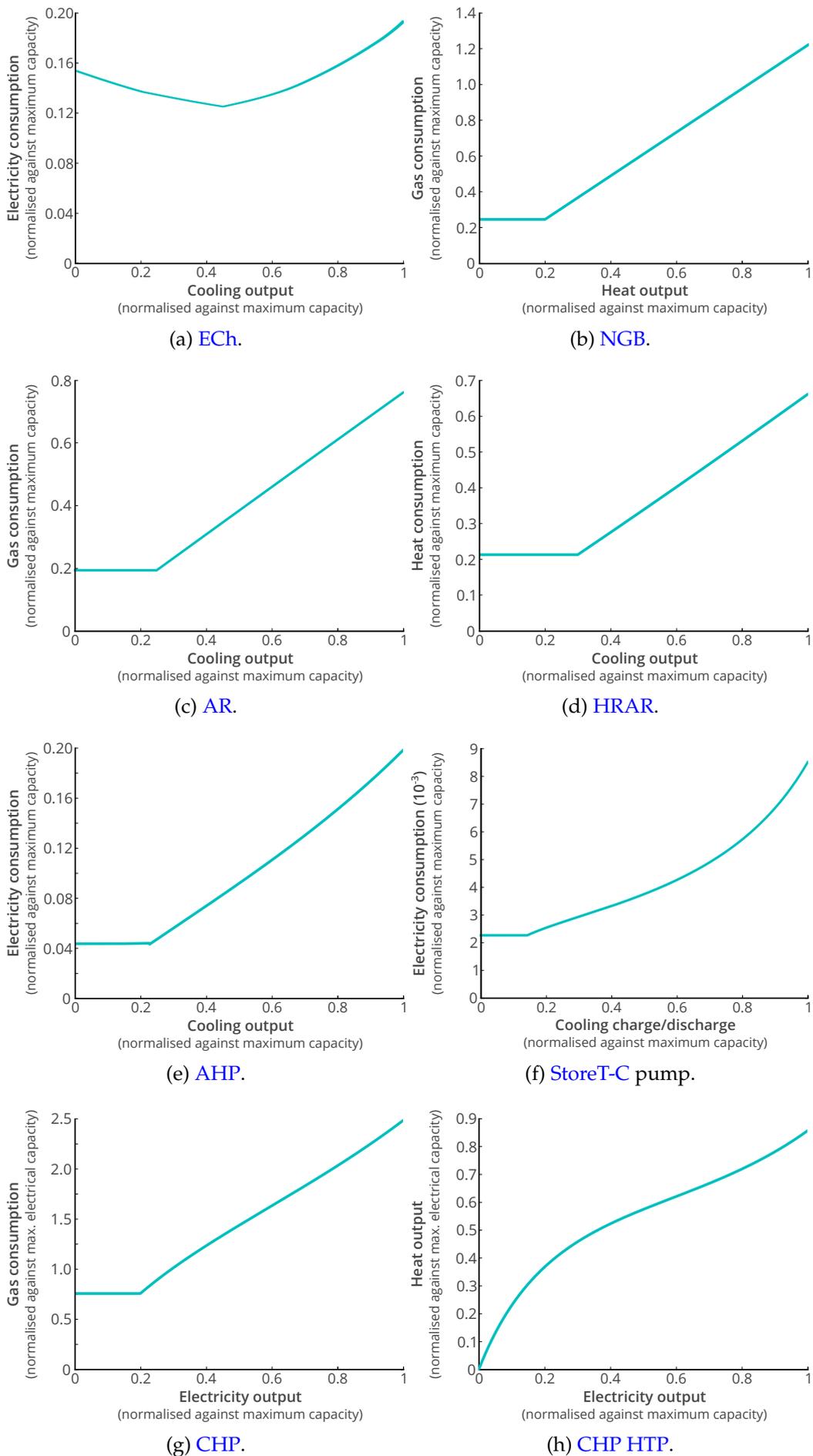


Figure 4.4 Nonlinear part-load curves for various energy technologies. All figures show energy consumption, except 4.4h which shows the **CHP** heat output per unit electricity output.

Table 4.1 Supply technologies and their consumption/production energy. E = electricity, G = gas, S = solar radiation, C = cooling, H = heating.

Technology	AHP	ECh	HRAR	AR	CHP <sup>a</sup>	NGB	PV	ST	StoreE	StoreT-H	StoreT-C
Consumption	E	E	G/H	G	G	G	S	S	E	H	C
Production	C	C	C	C	E, H	H	E	H	E	H	C
Rated capacity	550kW/ 500kW <sup>b</sup>	2500kW	1000kW	1500kW	352kW <sub>e</sub> 300kW <sub>th</sub>	750kW	1.2kW/m <sup>2</sup>	N/A	1200kWh	10000kWh	500kWh
Rated efficiency	358%	603%	132%	125%	40.5% 34.5%	80.0%	85% <sup>c</sup>	100% <sup>c</sup>	90%	100%	100%
Minimum load rate	0.2	0.45	0.30	0.25	0.2	0.2	0	0	0	0	0
Rated charge/discharge					N/A				360kW	3000kW	100kW

<sup>a</sup> CHP produces both electricity and heat.

<sup>b</sup> Two AHPs were used in this study, with differing maximum capacity.

<sup>c</sup> Solar energy conversion efficiency has already been accounted for in the timeseries resource data for PV and ST. For PV, 0.85 refers to inverter efficiency.

profile alongside a knowledge of the size and electricity price of the network power stations, as formulated by Ikeda and Ooka (2016). To create this case study, data on energy demand, technology characteristic curves, and costs have been brought together from multiple sources. Energy demand specific to a hotel was taken from the The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE) (2003) database and Japanese Meteorological Agency data on ambient conditions was used for energy consumption modelling of cooling technologies. This case study utilises the rated capacity of technologies given in Table 4.1 as it is only used for optimisation of the operation schedule. A full description of this case study, including the optimisation problem size, is given in Section B.1.

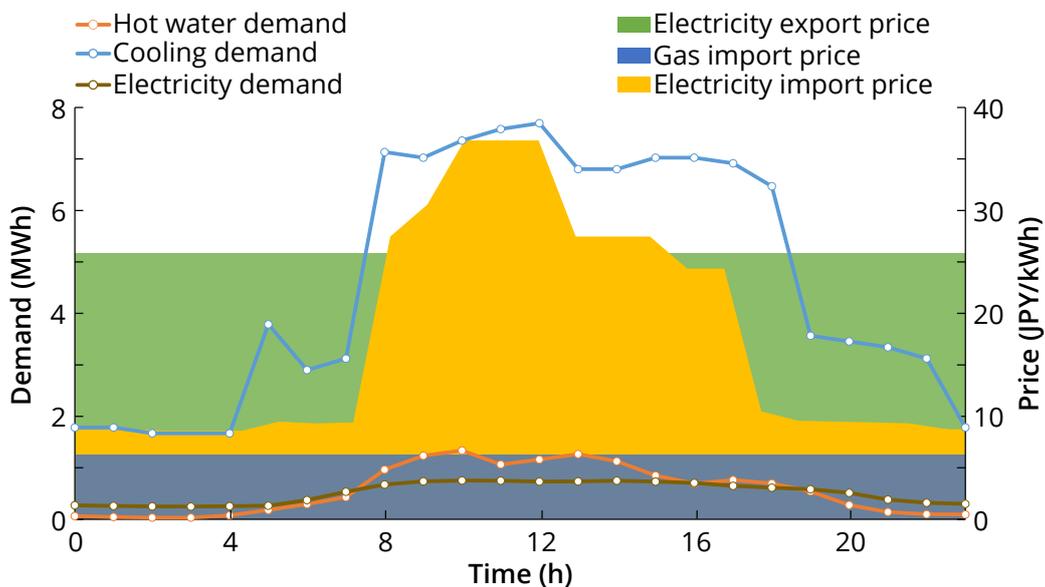


Figure 4.5 Energy demand and fuel pricing for case study 1, according to The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE) (2003) and Ikeda and Ooka (2016).

## Case study 2: UK district

Extending the single building system of case study 1, a district is considered in this case study. A hotel is a useful study building, due to its non-negligible requirement for electricity,

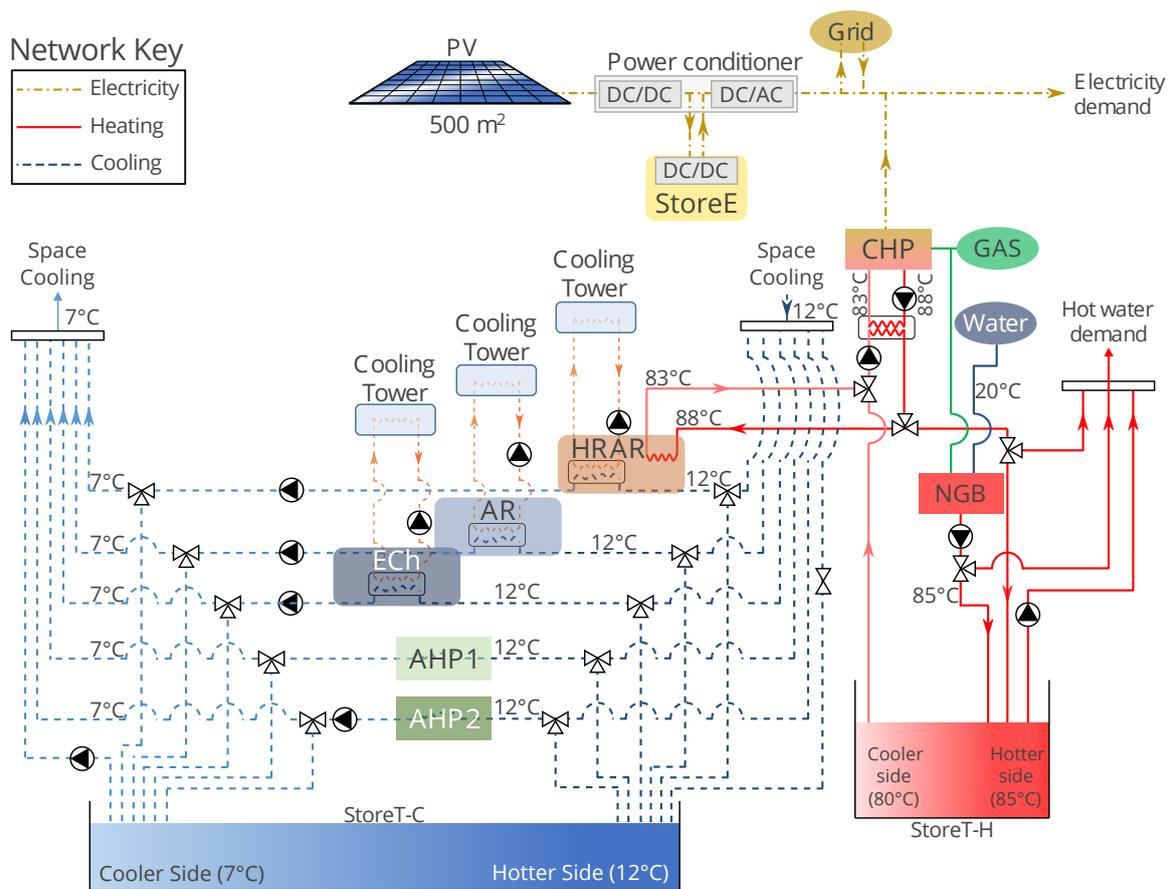


Figure 4.6 Schematic of supply and storage technologies for meeting multi-energy demand in case study 1. For clarification on abbreviations, see the list of acronyms.

heating, and cooling. But, unlike a district system, it does not require the consideration of energy distribution. This district is illustrative and consists of 10 domestic properties, one large hotel, one large office, and one power plant (Figure 4.7a). Different technologies are available in each building, depending on building demand and category. For instance, if a CHP is chosen, it must be situated in the power plant or in the hotel (as a micro CHP ( $\mu$ CHP)). No cooling technologies are considered in the domestic properties as there is only electricity and heat demand at those nodes. Table 4.2 gives further information on attributes of each property type. Distribution networks exist for low voltage electricity, gas, and heat. For tractability, the district is modelled in an aggregated form, as shown in Figure 4.7b. Commercial and domestic properties are aggregated into separate nodes in the district, with the buildings within each category contributing to its node. The district is located in the South-East of England, UK. However, due to limited data availability, U.S. Department of Energy representative building demand (EERE, 2013) informed hourly heat, cooling, and electricity demand of district buildings. Seattle, WA data was chosen due to climate similarity with London, UK. Demand data for an entire year was acquired, but was aggregated to four typical days (96 timesteps) when comparing piecewise to other linearised consumption curves. Two separate weeks (winter & summer, 168 timesteps each) were modelled when analysing equidistant against optimised breakpoint allocation.

Unlike case study 1, capacity and timeseries decision variables are considered in this district. Timeseries costs are those incurred by the consumption of national grid electricity (GridE) (0.095 GBP/kWh) and national grid natural gas (GridNG) (0.025 GBP/kWh). Revenue is available by the export of electricity from both the PV and CHP technologies. PV revenue is based on the feed-in tariff for domestic/commercial properties: 0.1203 GBP/kWh produced and 0.0491 GBP/kWh exported, respectively (ofgem, 2016). The CHP can export electricity at a price equal to 80% of the wholesale electricity price at any given time interval, using hourly wholesale prices for 2015, courtesy of ELEXON (<https://www.elexon.co.uk/>). Investment costs for each technology are given in Table 4.3. Most costs have been calculated based on values given in the SPON’S mechanical and electrical services price book (AECOM, 2015). These costs include a fixed installation cost and a cost that increases linearly with technology capacity. StoreT, StoreE, PV, and ST costs have been aggregated from online suppliers. A full description of this case study, including the optimisation problem size, is given in Section B.2.

Table 4.2 Case study 2 building characteristics.

	Dwelling	Hotel	Office	Plant	
Annual energy demand (MWh)	<i>Electricity</i>	7.2	1595.5	481.3	0
	<i>Heat</i>	17.5	1641.6	86.5	0
	<i>Cooling</i>	0.0	1757.9	99.1	0
Available roof area (m <sup>2</sup> )	130	1300	900	0	
Available technologies	NGB, PV, ST, StoreE, StoreT	$\mu$ CHP, NGB, PV, HRAR, AHP, ECh, ST, StoreE, StoreT	NGB, PV, ST, HRAR, AHP, ECh, StoreE, StoreT	CHP, GridE	

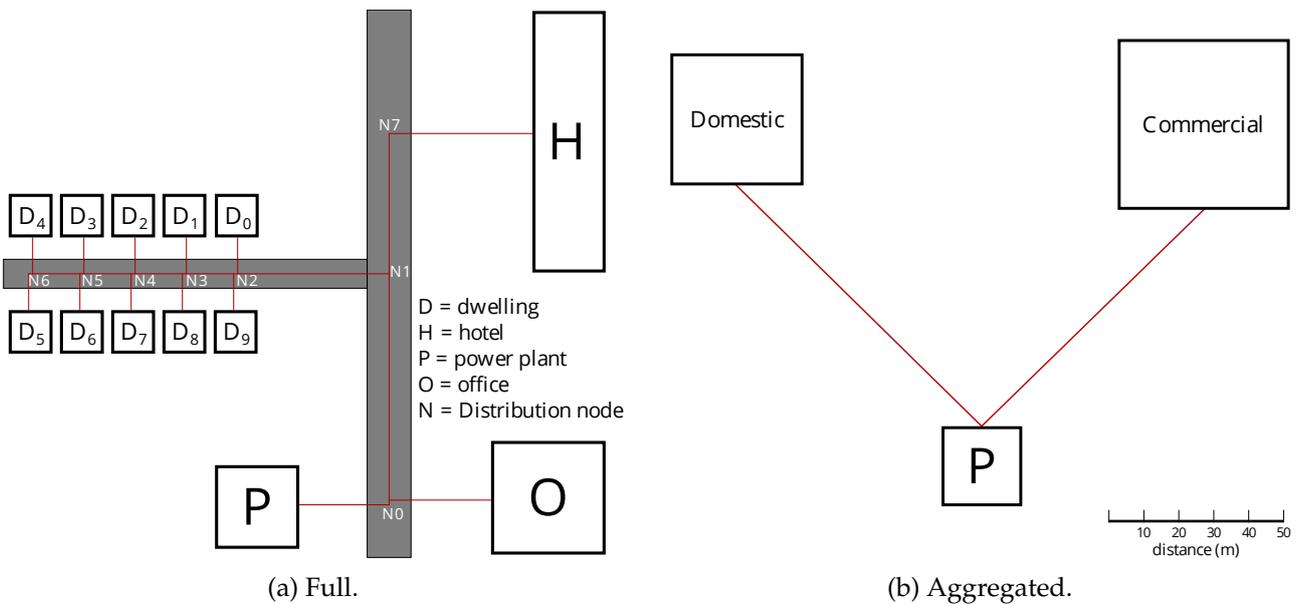


Figure 4.7 Graphical representation of case study 2 district network.

Table 4.3 Capacity costs (GBP) of supply technologies in case study 2.

Technology	AHP	ECh	HRAR	CHP	NGB	PV	ST	StoreE	StoreT-H	StoreT-C
Investment cost	2,517 + 158/kW	111/kW	52,497 + 79/kW	46,480 + 703/kW	2,024 + 35/kW	1,500 + 1,000/kW	1200/kW	1,667 + 350/kWh	527 + 66/kWh	527 + 303/kWh

### 4.2.3 Tests

In exploring the requirement for, and capabilities of linearisation, the case studies were used for two sets of tests. Various runs were undertaken for each test, as detailed in Table 4.4. The nonlinear curves are progressively linearised until they reach rated efficiency. In between, they are described by piecewise curves and  $y$ -intercept efficiency. This progression is illustrated in Figure 4.8. The implementation of  $y$ -intercept efficiency is one such that the efficiency at the minimum load of a technology is equal to the efficiency described by the nonlinear curve at that point. As such, technologies were not allowed to operate below their minimum load in these runs. There are a few instances where there is no cross-over between methods, such that a direct comparison cannot be made. Nevertheless, the impact of these methods relative to straight line efficiency and full nonlinear part-load curves allows for viable indirect comparison.

Breakpoint allocation can be purely subjective, involving choosing locations along the curve to define the minimum load rate; load rate at rated efficiency; and maximum load rate, with remaining points placed to ensure reasonable curve following. Using the second case study, methods of positioning the breakpoints along the curve, equidistant and optimised, and increasing the numbers of breakpoints are tested. Equidistant breakpoints are placed at regular intervals along the x-axis (load rate) of the curves. Optimised breakpoints are positioned using the method introduced in Section 4.1.

Table 4.4 Run configuration options for comparing nonlinear curves to their SLE and piecewise linearised counterparts. ‘x’ refers to the particular curve representation method available in each test.

	Representation method				Case study	Test
	SLE		Piecewise			
	Rated	y-intercept	CBS	SOS2		
Nonlinear vs piecewise	x		x		1	1
Nonlinear vs SLE	x			x	1	1
Piecewise vs SLE (1)	x		x		1	1
Piecewise vs SLE (2)	x	x		x	2	1
Piecewise varying breakpoints	x		x		2	2

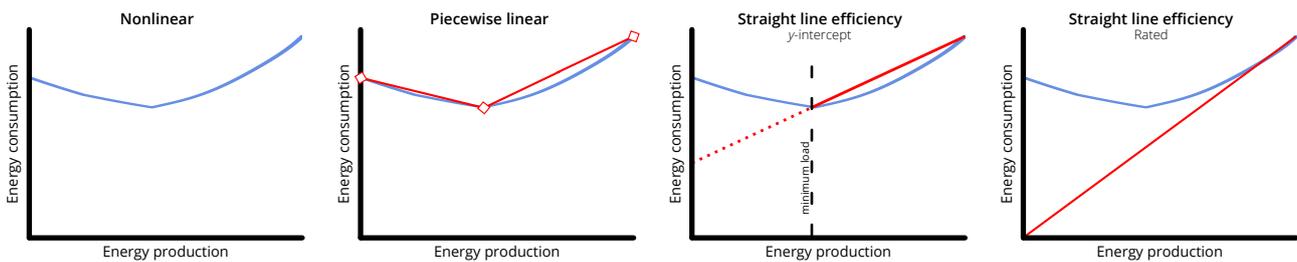


Figure 4.8 Progression of curve linearisation used in tests, from nonlinear to straight line.

### Nonlinear algorithm

In the first test, the impact of piecewise linearisation of part-load curves is explored and compared directly to their nonlinear counterparts. To this end, a nonlinear algorithm is required. As discussed in Section 2.2.1, metaheuristic algorithms are commonly used in energy system optimisation to handle nonlinearities. However, such algorithms are only used on small problems, due to the inefficiencies exhibited in trial and error optimisation. Ordinarily speaking, metaheuristic algorithms do not consider constraints, although almost all actual systems have many constraints. Thus, a method that can handle constraints is necessary. [Epsilon constraint differential evolution \( \$\epsilon\$ DE\)](#), developed by Takahama and Sakai (2010) and Mallipeddi *et al.* (2012), is one such method which exhibits efficiency improvements compared to other constraint handling methods. The  $\epsilon$ DE algorithm used in this study is that introduced by Ikeda and Ooka (2016) and contains seven parts. It is run in MATLAB (The Mathworks, Inc., 2015) as follows:

1. scatter 80 individuals into a search domain using random numbers;
2. evaluate the objective function of each individual and set an initial value of  $\epsilon$ ;
3. consider individuals only from the top 20% of the population, ranked by  $\epsilon$  level comparison;
4. create a new solution at a constant mutation rate of 0.5, mixing three individuals in accordance with the original differential evolution algorithm;
5. evaluate the constraint violation that indicates how far the new individual is from a feasible domain;

6. compare the new individual with the old in terms of the objective function or  $\epsilon$ , the better of the two becomes the old individual in the next iteration;
7. finally, return to 4, repeating generations up to 5000 times with cross-over rate decreasing exponentially from 0.8 and  $\epsilon$  having to reach zero at 2000 generations.

The most favourable objective function value from running the algorithm 10 times is taken as the optimal solution. Metaheuristic algorithms are often time consuming, particularly when solutions close to the global optimum are sought. The first case study is a building-level energy system to ensure timely convergence of the  $\epsilon$ DE model. The small spatial and temporal size of the problem allows technology nonlinear characteristic curves and the impact of temperature dependence on those curves to be examined. It also allows 5-breakpoint SOS2 to be used to describe the piecewise linearised curves.

**Temperature dependence** The ECh, AR and HRAR utilise a cooling tower to remove heat to the atmosphere. The performance of the cooling tower varies depending on external temperature, which can then be translated to lower average cooling water temperature ( $T_{cw}$ ). Lower  $T_{cw}$  leads to a lower average temperature of the respective technology refrigeration cycle, increasing its efficiency. This feedback loop can be modelled by metaheuristic algorithms, as temperature can be tracked within the primary optimisation. It is not feasible to emulate this behaviour in MILP, therefore the  $\epsilon$ DE model must be simplified to not account for  $T_{cw}$  changes. The effect of omission on the final objective function result will be analysed.

### Software and Hardware

When comparing solution times, it is important to ensure that there is consistency across both software and hardware in tests. Case study 1 tests were run on a laptop computer with 64-bit Windows 8.1 Pro operating system with 3.40 GHz i7-4700 processor and 32 GB RAM. Case study 2 tests were run on a 64-bit Windows 7 operating system with 2.50 GHz Intel Xeon E5-2680 v3 processor and 64GB RAM. All MILP models were optimised using IBM ILOG CPLEX (IBM Corp., 2016), with a variation of the modelling framework DENO (Omu *et al.*, 2013; Omu *et al.*, 2015) used in case study 1 and the Calliope framework (see Section 3.2) used in case study 2.

## 4.3 Results

### 4.3.1 Influence of linearising technology consumption curves

#### Objective function value

In case study 1, Table 4.5 shows that the OFV range is bounded by the simplest ( $y$ -intercept efficiency) and most complex ( $\epsilon$ DE with  $\Delta T_{cw}$ ) curve representation. The difference of 3% is small and initially the piecewise case provides a better solution than using  $\epsilon$ DE. This clearly indicates that  $\epsilon$ DE did not reach a global optimum, as piecewise characteristic curves are more pessimistic than the nonlinear curves in every case. That is, the optimum that could

be achieved by  $\epsilon$ DE is lower than in the MILP models, for the same operation schedule. The  $\epsilon$ DE result is brought marginally lower than the piecewise one by updating the method given in Section 4.2.3 to initialise with MILP piecewise results, instead of randomly scattering individuals. By doing this, it was possible to obtain a marginally lower operation cost, which just beats the piecewise result.

When accounting for linearisation error (LE) ex-post, the piecewise result was 1,500 JPY lower than given by its OFV. Better yet, the operation cost when using  $\gamma$ -intercept efficiency decreased by 5,400 JPY ex-post. Thus, the operation cost based on linear models is lower than that achieved using  $\epsilon$ DE, even when initialising it with the piecewise results. A greater improvement can be made by allowing temperature tracking in  $\epsilon$ DE, as the cooling technologies operate to capitalise on periods of high COP. But, given the inability of  $\epsilon$ DE to find a global optimum without varying  $\Delta T_{cw}$ , it is likely there is an even lower cost operating schedule that has yet to be found.

Table 4.5 Objective function value and optimisation solution time for case study 1, comparing linear, piecewise linear, and nonlinear technology representation. LE is calculated ex-post.

Run	MILP		$\epsilon$ DE		
	$\gamma$ -intercept efficiency	Piecewise linear	Random initialisation	MILP initialisation	Random initialisation inc. $\Delta T_{cw}$
OFV (JPY)	710,542	705,292	706,013	705,050	686,527
OFV + LE	705,150	703,895	706,013	705,050	686,527
Solution time (s)	0.08	1.64	276	270	295

For the district system in case study 2, Figure 4.6 shows that the result is varied. If using the  $\gamma$ -intercept efficiency, as used in case study 1, the OFV is slightly lower than piecewise. However, after accounting for LE ex-post, the cost is actually higher. If using piecewise linearisation, the ex-post error is low: actual operation will cost 0.1% less than the OFV. Once LE has been applied to rated efficiency runs, the operation cost increases by 5%, becoming the most expensive option.

Table 4.6 Objective function value and optimisation solution time for case study 2, comparing linear and piecewise linear technology representation. LE is calculated ex-post. When modelling piecewise curves, three breakpoints were used.

Run	Rated efficiency	$\gamma$ -intercept efficiency	Piecewise linear
OFV (GBP)	227,332	232,801	233,173
OFV + LE	239,352	232,630	231,714
Solution time (s)	186	401	598

### Solution time

By increasing model complexity, time to solution increases. There is a two order of magnitude time saving when using  $y$ -intercept efficiency compared to piecewise representation (SOS2) and a further two orders of magnitude compared to the  $\epsilon$ DE solution time. The  $\epsilon$ DE model operation varies little between each run type, as the number of initial individuals and generations to run within the algorithm are fixed across all runs. The 7% increase in time for  $T_{cw}$  consideration shows the computational effort within MATLAB to calculate updated temperatures at each generation. Compared to SOS2,  $y$ -intercept efficiency is fast. Compared to CBS,  $y$ -intercept efficiency is (slightly) slow. Case study 2 shows that CBS is a drastic improvement on SOS2 as its solution time is of the same order of magnitude as  $y$ -intercept efficiency. The fastest solution time is realised when using rated efficiency, owing to the model requiring the least number of binary variables.

### Capacity and dispatch

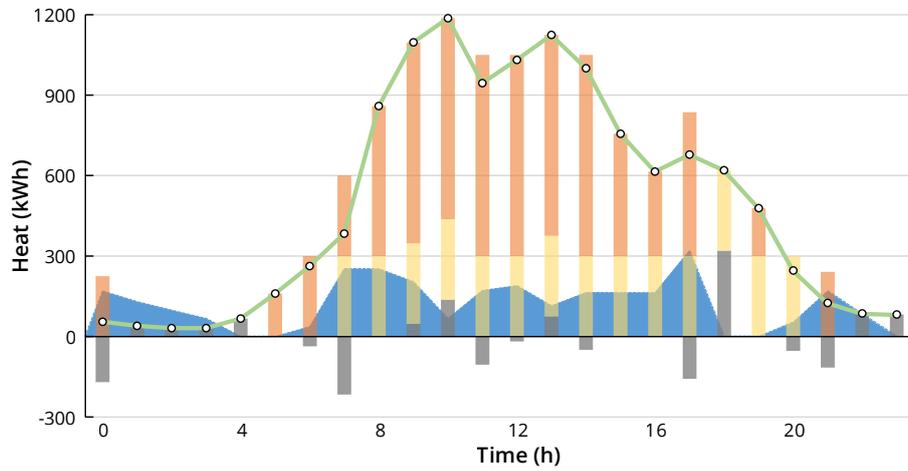
Across the 24 hours of operation, the energy dispatch can be readily compared in case study 1. Only minor differences exist between the models, but comparison of the  $\epsilon$ DE and MILP results shows different functions for the hot water storage tank (Figure 4.9). In the  $\epsilon$ DE local optimal case, the early hours are used to charge the heat storage to 600kWh, which is then discharged beginning at 6am. Storage is more sporadic in the MILP case, never charging the heat storage to more than 300kWh, and more often changing between charge/discharge from hour to hour.

There is no discernible difference in operation between the two linear representations of case study 1. In case study 2, Figure 4.10 shows clear differences in both technology investment and in operation. First, there is far more dependence on the ECh when there is no knowledge of its part-load operation (i.e. using rated efficiency) as it can be used liberally at part-load without any negative impact. Full load operation of a smaller capacity ECh is used in both the  $y$ -intercept and piecewise schedules. Having more technologies with lower capacities either at full load or zero load is particularly prevalent in the piecewise case, which is the only solution to invest in AHP. Capacity of heating and storage technologies is uniform across each of the runs. Although there are differences in the operation of the NGB and CHP, there is no clearly discernible trend with greater simplicity of the linearisation method.

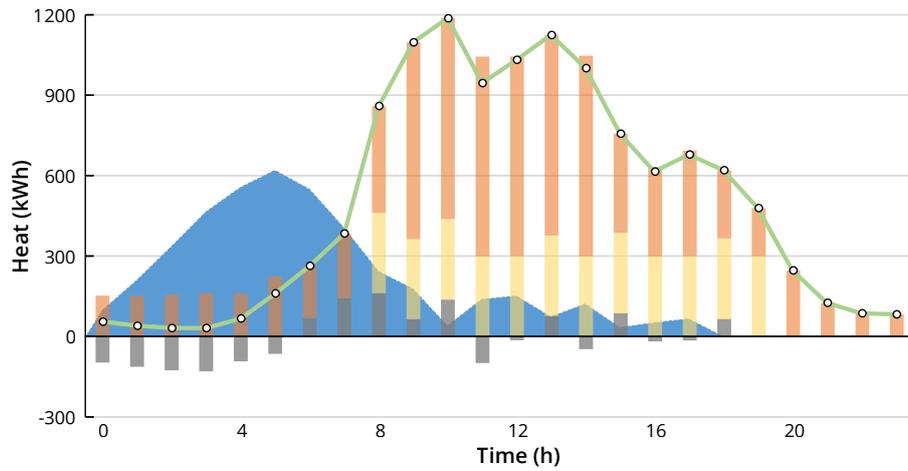
## 4.3.2 Breakpoint allocation

### Ex-ante linearisation error

Figure 4.11 shows the difference in breakpoint positioning between the optimised and equidistant allocation methods. Equidistant breakpoints deviate around the minimum load, which is avoided when optimising the breakpoints. Also, with an increasing number of breakpoints, the ex-ante LE reduces much more quickly with optimised piecewise curves than with equidistant ones. Realistically, there is only a small variation between the two sets of curves, but the advantage of the optimised breakpoint allocation is that no prior knowledge of the curve is required. The optimisation process took only 17.1 seconds to optimise 108 piecewise curves describing characteristics of eight technologies (27 nonlinear



(a) MILP piecewise linear.



(b)  $\epsilon$ DE.

● Hot water demand    
 ■ NGB output    
 ■ CHP output    
 ■ Stored Heat    
 ■ Store-T-H charge/discharge

Figure 4.9 Comparison of optimal operation schedules for meeting hot water demand in case study 1.

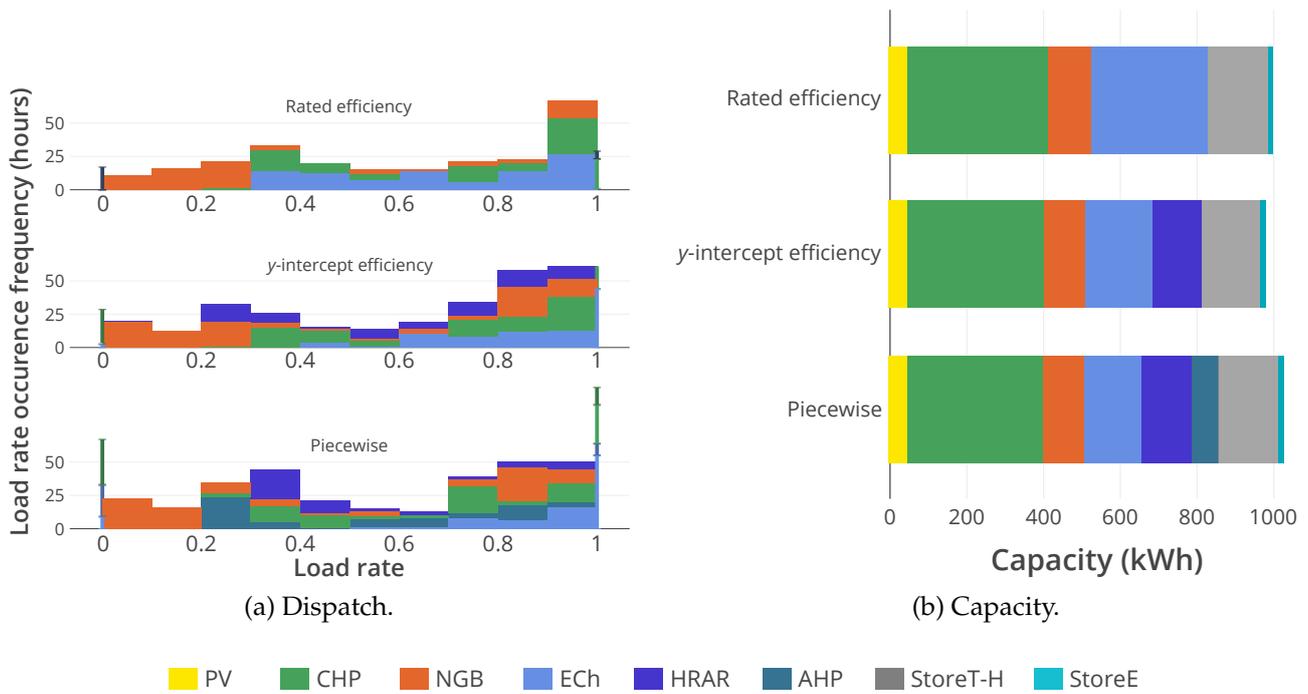
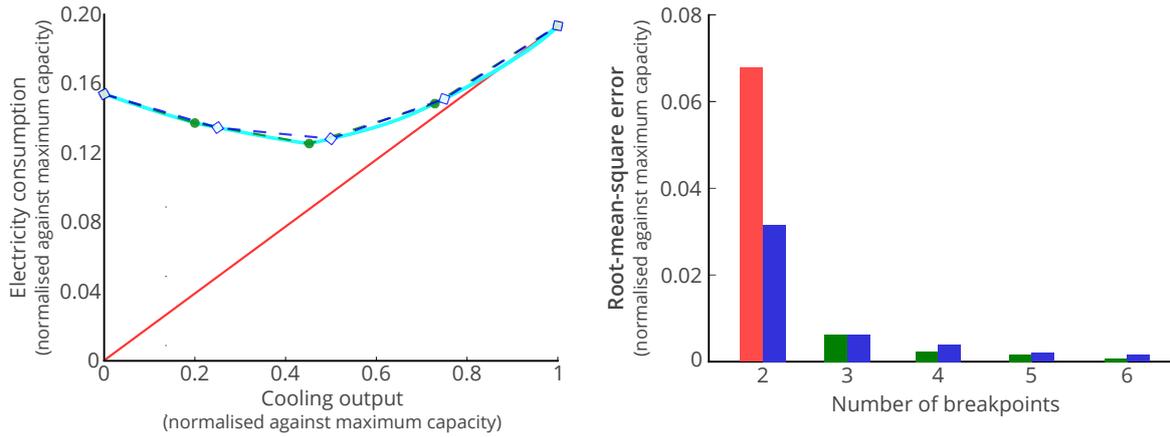


Figure 4.10 Comparison of optimal capacity and dispatch between different part-load characteristic linearisation methods in case study 2. When modelling piecewise curves, three breakpoints were used. Dispatch shows a histogram of the number of hours a technology operates at a given load rate, with zero and full load given at the extreme ends of the histograms.

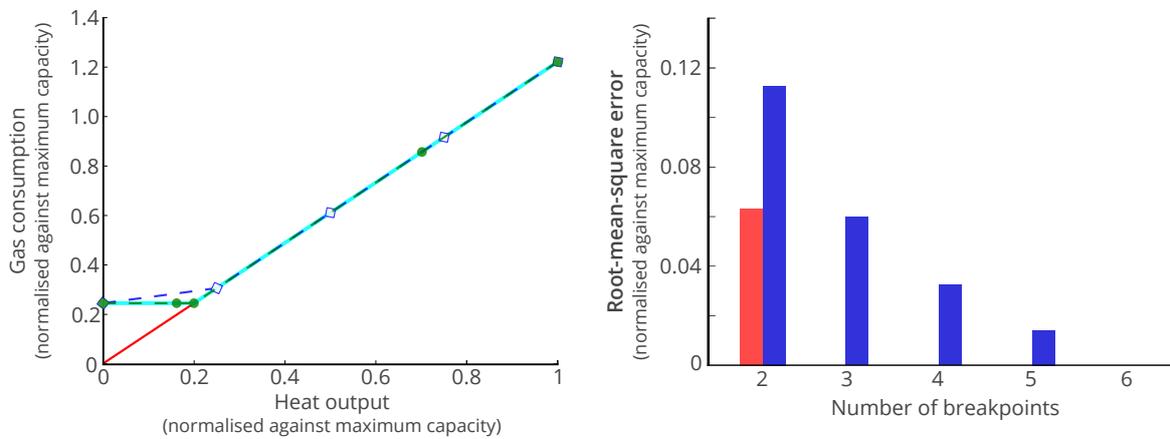
curves, three to six breakpoints). This optimisation speed would not differ greatly if more complex curves were applied, but equidistant breakpoints may become more erroneous. Additionally, the **CHP** part-load curves (Figures 4.11e and 4.11f) show the advantage of an optimisation technique to position breakpoints: a constraint to ensure **CBS** can be enforced. The additional error in ensuring that the piecewise curve has a strictly increasing/decreasing gradient is apparent at a higher number of breakpoints. But, in all but one instance, the **RMSE** is still lower than equidistant breakpoint allocation, with the added benefit of knowing that **CBS** can be used.

### Objective function value

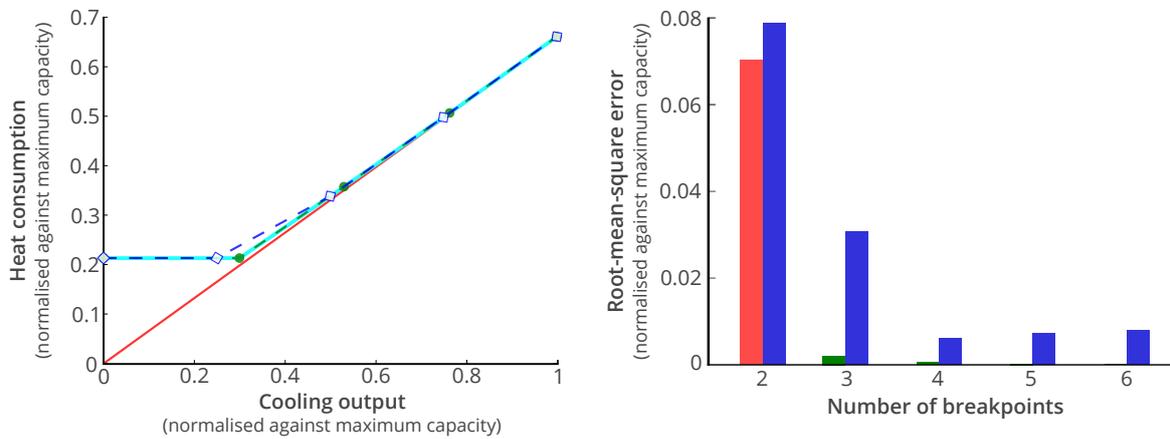
Application of piecewise curves increases the objective function value by as much as 5.2%. Table 4.7 shows that differences in objective function value are small when increasing the number of piecewise breakpoints, with optimised curve averages of £4036 +1%/-0.5% in winter and -£2394 +0%/-0.6% in summer. The summer negative cost represents the ability for the system to gain more revenue from subsidies and export than it spends on investment and operation in that period. There are no equidistant solutions beyond three breakpoints due to model infeasibility. It is not possible to place constraints on breakpoint location when placing equidistantly. Thus, the strictly increasing/decreasing gradient requirement for being bound by constraints cannot be met for **CHP** HTP and gas consumption. Although the rated efficiency objective function value is lower than for piecewise models, the 'real' system costs end up being higher. Ex-post **LE** is 12% in both seasonal weeks when using rated efficiency,



(a) ECh.



(b) NGB.



(c) HRAR.

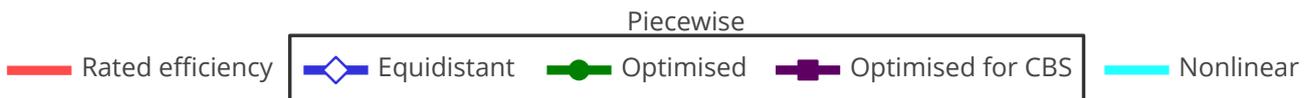
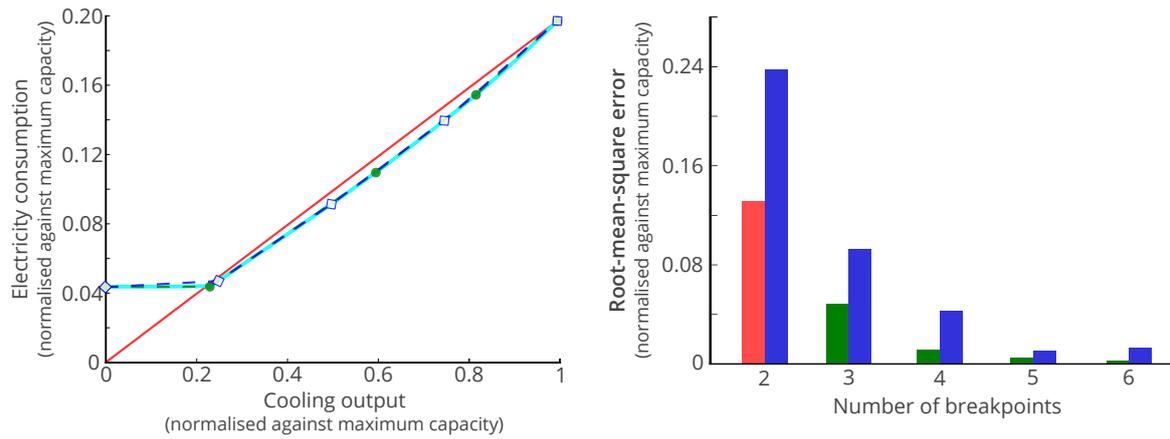
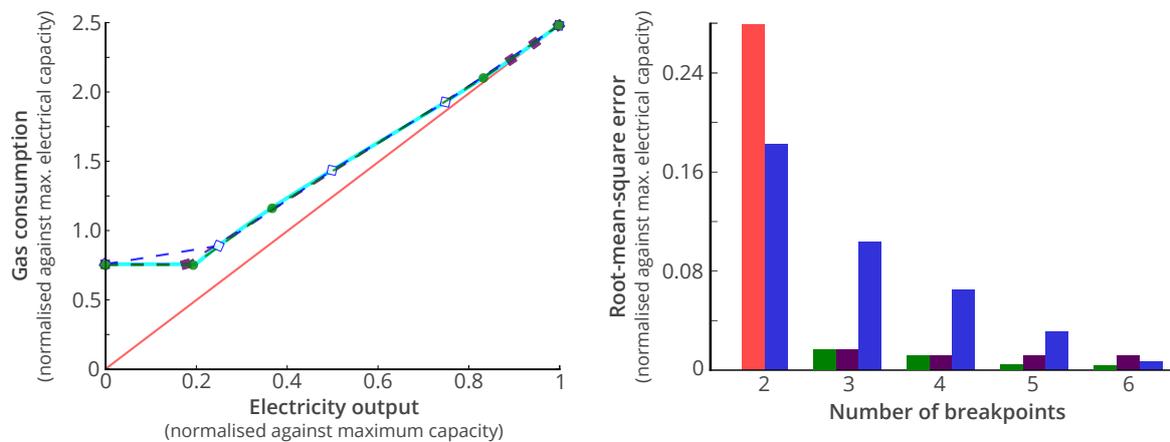


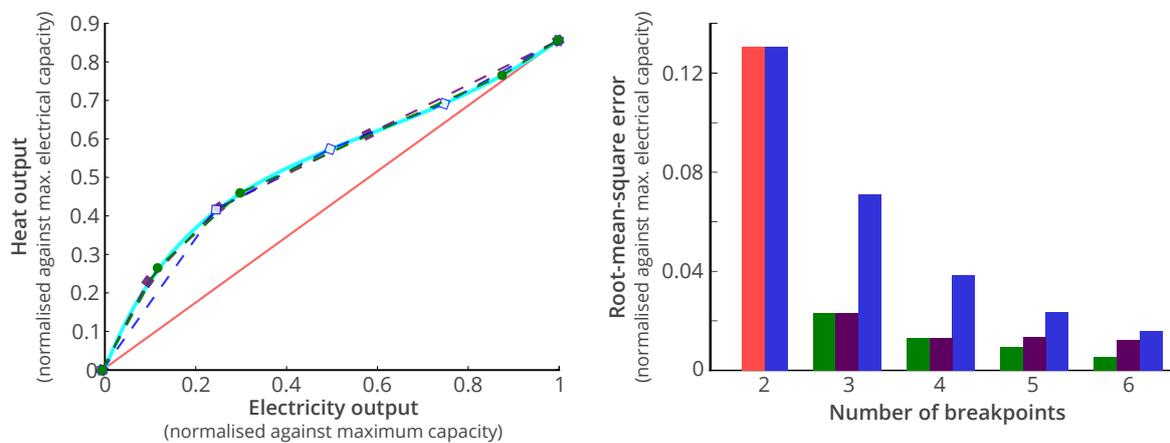
Figure 4.11 Impact of optimised and equidistant piecewise curves on ex-ante LE of part-load curves of technologies considered in case study 2. 5-breakpoint piecewise and rated efficiency curves shown for comparison against nonlinear curve. All figures have energy consumption on the y-axis, except 4.11f which shows the CHP heat output per unit electricity output. The 'optimised for CBS' results for the CHP refer to the requirement that the piecewise curve has a strictly increasing/decreasing gradient.



(d) AHP.



(e) CHP.



(f) CHP HTP.

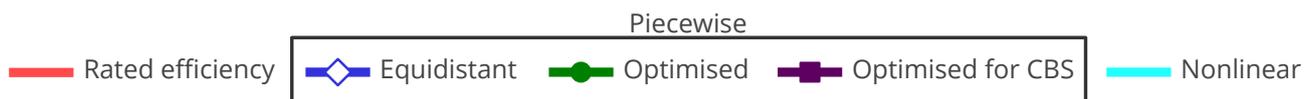


Figure 4.11 (cont.) Impact of optimised and equidistant piecewise curves on ex-ante LE of part-load curves of technologies considered in case study 2. 5-breakpoint piecewise and rated efficiency curves shown for comparison against nonlinear curve. All figures have energy consumption on the y-axis, except 4.11f which shows the CHP heat output per unit electricity output. The 'optimised for CBS' results for the CHP refer to the requirement that the piecewise curve has a strictly increasing/decreasing gradient.

decreasing to less than 1% when including piecewise curves. In summer, this effect is most pronounced, where LE reduces to zero at six breakpoints.

Table 4.7 OFV (in GBP) for all breakpoint comparison run configurations, including the quantification of ex-post LE. O = optimised, E = equidistant, RE = rated efficiency.

<i>Breakpoints</i>	2	3		4		5		6	
<i>Linearisation</i>	RE	O	E	O	E	O	E	O	E
<b>Winter</b>									
<b>OFV</b>	3,989	4,036	4,048	4,074	Fail	4,019	Fail	4,016	Fail
<b>+LE</b>	+465	+23	+12	+40	N/A	+32	N/A	+32	N/A
<b>Summer</b>									
<b>OFV</b>	-2,507	-2,380	-2,377	-2,398	Fail	-2,398	Fail	-2,401	Fail
<b>+LE</b>	+294	-17	-28	-2	N/A	-1	N/A	0	N/A

### Solution time

While the accuracy of the objective function value is improved, piecewise linearised cases take much longer to solve (Table 4.8). This is more the case in the summer week, which peaks at 17,521 seconds (three breakpoints, equidistant), two orders of magnitude greater than the rated efficiency run. Even at the least number of breakpoints, the solution time is 2.5x and 14.9x longer than the rated efficiency run in winter and summer, respectively.

There is generally an increase in solution time with an increased number of breakpoints, the only anomaly being the drastic decrease in model solution time between having five and six breakpoints in summer. Here, the optimisation is completed in less than half the time with an additional breakpoint. In this instance, the five-breakpoint case had been solved within 10% of the relaxed LP 200 seconds sooner than the six-breakpoint case, but failed to converge on the last few percent for an extended period. Equidistant breakpoints decrease the solution time by a small amount in the winter week and increase it substantially in the summer week. Again, it is the final few percent of convergence that leads to the vastly inflated solution time.

Table 4.8 Model runtime in seconds for all configurations, including pre-processing and subsequent optimisation. O = optimised, E = equidistant, RE = rated efficiency.

<i>No. of breakpoints</i>	2	3		4		5		6	
<i>Linearisation</i>	RE	O	E	O	E	O	E	O	E
<b>Winter</b>	366	926	610	880	Fail	847	Fail	1,408	Fail
<b>Summer</b>	300	4,483	17,521	7,202	Fail	15,230	Fail	6,816	Fail

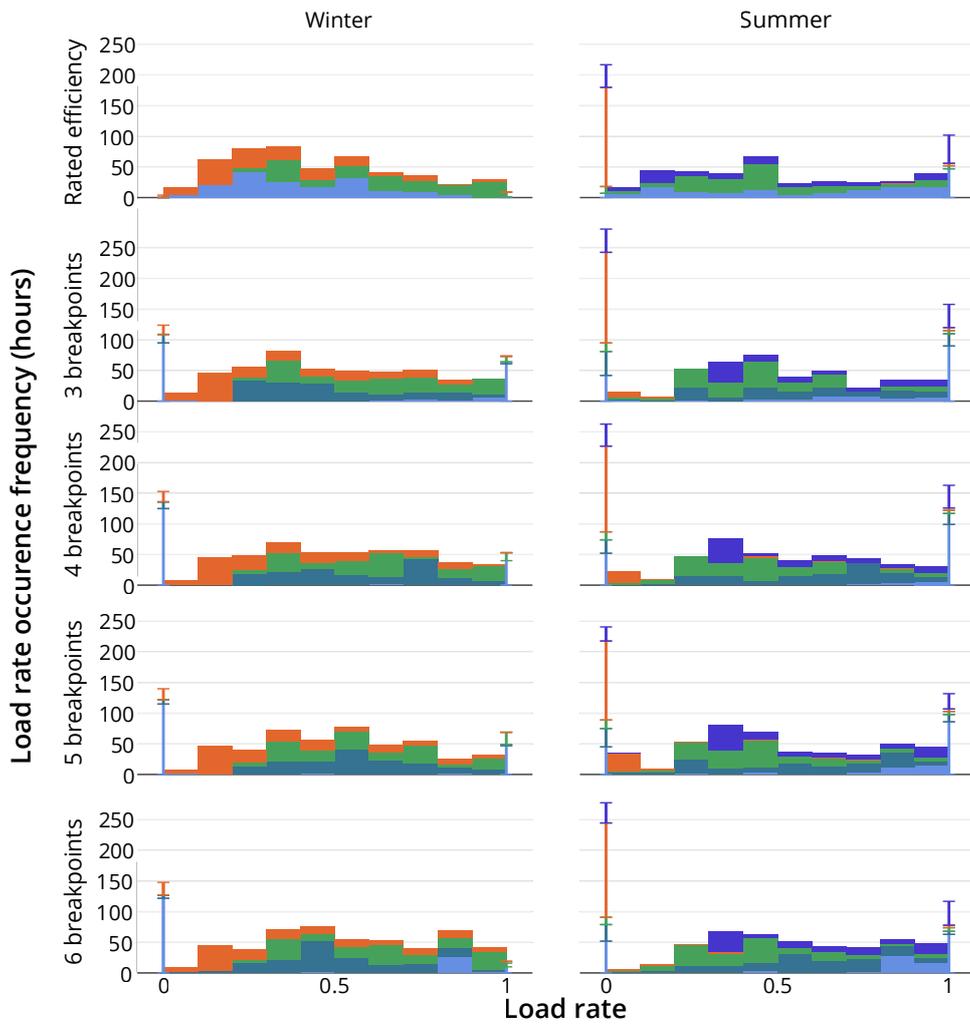
### Capacity and dispatch

The change of objective function value when applying piecewise characteristic curves results from changes in both investment and operation. Varying the impact of part-load operation leads to different technology choices. For instance, in meeting cooling demand

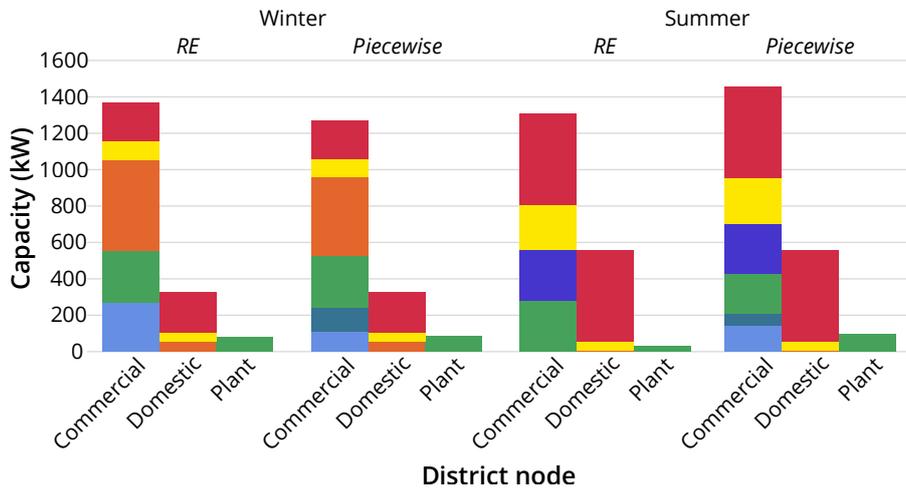
in the rated efficiency run, the **ECh** is chosen to operate as the only technology throughout. When applying piecewise curves, Figure 4.12a shows that **AHP** is better suited for part-load requirements, leaving the **ECh** for almost exclusive use at its full load. Generally, there is more use of technologies in full/zero load configurations when piecewise curves are included. This means that a greater variety of technologies are purchased to avoid running any one of them at part-load. Purchased technology capacities also vary (Figure 4.12b). In both seasons, **ECh** capacity is reduced in the piecewise results and **AHP** is purchased to account for the deficit. In the winter week, **NGB** size is also reduced, balanced by a larger heat storage capacity (Table 4.9). Storage is used more in piecewise models, leading to lower cumulative system capacity. These results also show that the utility of the local distribution network is dictated by technology choices. For example, more power is distributed to the commercial properties in summer due to the purchase of a smaller  $\mu$ **CHP**. Although a small **CHP** is purchased in all cases, a heat network is avoided. Instead, heat is dumped to the atmosphere. However, the system is limited in how much heat it can dump, so the plant **CHP** could feasibly be larger if that constraint were lifted.

Table 4.9 Comparison between piecewise (P) and rated efficiency (RE) results of distribution network between, and storage capacity at, demand locations of case study 2. Only 3-breakpoint piecewise given, for clarity, since investment portfolios of piecewise models are all very similar.

	<i>Distribution</i>						<i>Storage</i>					
	Gas		Heat		Power		Cooling		Power		Heat	
	<i>RE</i>	<i>P</i>	<i>RE</i>	<i>P</i>	<i>RE</i>	<i>P</i>	<i>RE</i>	<i>P</i>	<i>RE</i>	<i>P</i>	<i>RE</i>	<i>P</i>
<b>Winter</b>												
commercial	<b>1304</b>	1224	0	0	71	<b>75</b>	0	<b>24</b>	7	0	0	<b>59</b>
domestic	<b>69</b>	67	0	0	<b>41</b>	37	0	0	7	7	145	145
<b>Summer</b>												
commercial	<b>788</b>	622	0	0	40	<b>135</b>	0	7	7	7	230	<b>289</b>
domestic	6	<b>9</b>	8	0	40	<b>43</b>	0	0	7	7	5	8



(a) Dispatch.



(b) Capacity.

EC AHP CHP NB HRAR PV ST

Figure 4.12 Comparison of optimal capacity and dispatch as the number of piecewise linear breakpoints is increased from two (rated efficiency) to six, for case study 2 operating in summer and winter. Dispatch shows a histogram of the number of hours a technology operates between 10% increments of load rate, with zero and full load given at the extreme ends of the histograms. Only 3-breakpoint piecewise given when comparing capacity, since investment portfolios of piecewise models are all very similar.

## 4.4 Discussion

When choosing how to represent technology part-load characteristics, piecewise linearisation has an influence on the objective function value. By better matching the nonlinear curves, the optimal system is more expensive in both case studies. This greater expense is more representative of the ‘real’ cost of the optimal system. In fact the ‘cheaper’ systems, as given when using nominal efficiency, are ultimately more expensive when the nonlinear curves are applied ex-post or the nonlinear models manage to reach their global optimum. But, the improvement in accuracy comes with a solution time penalty. Between studies, a range of four orders of magnitude in solution time is seen, depending on whether a nominal efficiency is used or nonlinearities are fully represented using  $\epsilon$ DE. The smallest jump is when using CBS to piecewise linearise, as there is at most an order of magnitude time penalty. When using SOS2 to piecewise linearise in the second case study, the time penalty compared to nominal efficiency is two orders of magnitude. Different solution times are acceptable for different schedule optimisation purposes. Although control systems which automate system management in real-time require instantaneous results from which to act, operation scheduling as part of system design is less time critical; a designer can wait for a result with no adverse effects. Timing requirements contrast with realistic representation requirements. Knowing that a battery must meet e.g. 30kW or 32kW at a given instant is more important five minutes before the instant than it is during the design phase. It is not currently possible to optimise with the most realistic, nonlinear representation of technologies in a time that is viable for instantaneous control. System designers may also find solution time prohibitive, especially when it invariably becomes several days for a larger problem, spatially and temporally. Design is iterative and to ensure that iterations can be checked it needs to be simple and quick.

If looking to use piecewise linearisation, it is possible to realise solution time improvements by both using CBS and optimising breakpoint placement. As an increasing number of breakpoints does not greatly reduce the curve fit error, it is possible to further limit the CBS piecewise time penalty by going no further than three breakpoints. The 2.5x time penalty in winter for introducing three breakpoints is probably justifiable. However, the 15x time penalty in summer may become unacceptable. Equally, choosing three equidistant piecewise breakpoints proves beneficial in the winter case (1.5x quicker than optimised breakpoints) but certainly not in the summer case (4x slower than optimised breakpoints). These three-breakpoint piecewise models have the same number of decision variables, which leads to agreement with Bischi *et al.* (2014) that breakpoint positioning is an important factor in piecewise linearisation. In equidistant positioning, the central breakpoint is at 50% load rate, which requires more branches to be searched, as there could be an optimal schedule with technologies operating either side. Conversely, the central breakpoint when optimising positioning, tends towards a 20-30% load rate. This breakpoint sits in a less critical part of the curve because so few possible solutions involve technologies operating at such low load rate. This problem of hopping either side of breakpoints is only exacerbated by a greater number of them, hence the solution time increasing with number of breakpoints. Thus, assigning

breakpoint locations to facilitate rapid convergence is difficult. Certainly, avoiding SOS2 and assigning breakpoints for strictly increasing/decreasing gradient is a good first step in reducing solution time, as the number of decision variables is reduced. Further study is required to better understand the critical nature of breakpoints. One possible method is to run the model with a lower tolerance (e.g. 10% gap between the MILP result and the LP relaxation). With this solution, the change in operation schedule will already be evident. Critical operating regions could then be identified, and breakpoints re-adjusted to avoid those regions. The design decision also depends on the purpose of the model: feasibility level studies could model with nominal efficiencies to get a lower system cost, then apply nonlinear curves ex-post to get an upper bound. The piecewise system cost will exist within that range. Only on undertaking detailed design, for investment portfolio and operating schedule, would piecewise curves be necessary.

With a particularly long time to prepare a model, it may become feasible to instead consider the nonlinear load curves, using an algorithm like  $\epsilon$ DE. Alongside a lengthy time to solution, the optimisation result of the  $\epsilon$ DE model displayed a completely different hot water operation schedule to that of the MILP model. As the operation cost varies so little, these scheduling differences provide an additional decision criterion of user subjectivity. On one hand, the MILP schedule would allow the user to downsize their hot water tank to half the size of that in the  $\epsilon$ DE case. But, the irregular charge/discharge could have adverse effects which are not bounded by the problem. For example, mixing hot/cold fluids in the stratified tank could have a noticeable impact on the stratification. Effects could also be beneficial, in which the greater quantity of heat stored during the day would lead to less heat loss, because ambient temperature is higher. Whether one schedule is more beneficial than another is difficult to discern, due to the complex thermodynamic processes. The model considers them to be equally good due to simplicity of storage representation across all algorithms.

Where load rates are used, the shape of the nonlinear curve clearly has an impact. The minimum load rate, which causes a rapid change in efficiency for most technologies, may vary between manufacturers. Certainly, the load curves used in this research do not entirely agree with those used elsewhere (e.g. Bischi *et al.*, 2014). When considering masterplan level design, there is little to no knowledge of the exact technologies that will be used, making it difficult to include nonlinear characteristics that will actually be representative of the technologies. This raises the question of whether piecewise linearisation of an incorrect curve is worse than using a nominal efficiency and, as such, whether it is worth the time penalty incurred in incorporating the perceived better representation.

## 4.5 Conclusions

This chapter has compared five representations of technology part-load consumption curves. They increase in simplicity, from the full nonlinear representation to the use of rated efficiency for all part-load conditions. In between, nonlinearities are described by linear pieces. All but the nonlinear representation can be incorporated into a [mixed integer linear programming \(MILP\)](#) model. Ultimately, [MILP](#) offers speed and simplicity. Piecewise linearisation allows gap bridging; it leads to slower solution times compared to using rated efficiency, but is fast compared to [epsilon constraint differential evolution \( \$\epsilon\$ DE\)](#). The operation schedule is also similar to  [\$\epsilon\$ DE](#), but lack of temperature tracking prevents it from being fully representative. The importance of temperature tracking would be further pronounced on a district level, where heated or cooled water is transported over long distances.

Within piecewise curve representations, further improvements in solution time are possible without loss of understanding. Most of the part-load curves in this chapter meet the requirements for representation using simple [convex bounding sets \(CBS\)](#), instead of the more complex [special order set constraint of type 2 \(SOS2\)](#), thanks to the strictly increasing gradient of part-load curves. Although the [combined heat and power plant \(CHP\)](#) part-load curves did not meet this requirement, automation of breakpoint allocation can be used to enforce it in the piecewise representation whilst minimising its error relative to the nonlinear curve. Indeed, automation proves to be a straightforward extension of piecewise linearisation methods. It can successfully be employed to select the best position of breakpoints along the piecewise curve, reducing the ex-ante error compared to heuristic allocation.

Although piecewise linearisation is a viable method to better represent technology characteristics, it may still prove to be too cumbersome in some analyses. Using [CBS](#) can lead to models which solve 15 times slower than when using a rated efficiency. It also requires that nonlinear characteristics for technologies are known. In this chapter, recommended representations from a government body were used, but their disagreement with other studies raises questions of the impact an incorrect curve would have on real operation of a system.



# Chapter 5

## Decision making under uncertainty

Parts of this chapter have been published as:

B. Pickering and R. Choudhary (Sept. 2018). ‘Mitigating Risk in District-Level Energy Investment Decisions by Scenario Optimisation’. en. In: *Proceedings of BSO 2018*. Cambridge, UK, pp. 38–45. URL: <http://www.ibpsa.org/proceedings/BSO2018/1B-1.pdf>

B. Pickering and R. Choudhary (Feb. 2019). ‘District Energy System Optimisation under Uncertain Demand: Handling Data-Driven Stochastic Profiles’. In: *Applied Energy* 236, pp. 1138–1157. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2018.12.037](https://doi.org/10.1016/j.apenergy.2018.12.037)

Although preparation of consumption data pertaining to the Bangalore case study was undertaken as part of this research, stochastically generated samples of demand were produced by Rebecca Ward, using [functional principal component analysis](#), as published in: R. Ward *et al.* (Sept. 2018). ‘A Stochastic Data-Centric Model for Quantification of End-Use Energy Demand in Buildings’. In: *Proceedings of BSO 2018*. Cambridge, UK, pp. 307–314. URL: <http://www.ibpsa.org/proceedings/BSO2018/3C-2.pdf> (visited on 23/10/2018).

Perfect foresight is the current paradigm in district energy system optimisation. A model created for planning purposes will consider a single instance of the future. In optimising for one possible realisation of parameter values, modellers ignore the inherent uncertainty exhibited by real systems. For instance, in simulating demand for UK commercial properties, the deviation between expected and metered annual demand has been reported to be anywhere between 16% and 500% (The Carbon Trust, [2011](#)). Therefore, if optimisation takes place using simulated demand, unexpected variations could result in system failure once commissioned.

Furthermore, [national grid electricity \(GridE\)](#) is assumed to be permanently available in district energy modelling and is a source of energy that is often chosen in preference to other technologies. In sensitivity analyses, Wouters *et al.* ([2015](#)) increased the cost of [GridE](#) in their small district, showing that even if the cost doubles, grid electricity is still the key electricity source in the system. Yet, although an average UK customer experiences fewer than 100 minutes of power interruptions per year (EQS WS, [2018](#)), the World Bank’s Enterprise Surveys (<http://www.enterprisesurveys.org>) show that 10% of electricity is supplied by auxiliary, backup generators in European firms, and rises to 46% in India.

The reality of system design is that safety factors are put in place to ensure that the installed capacity of a system is sufficiently large to meet unexpectedly high demand or to

overcome periods of power interruption. This approach could be considered as risk-unaware or particularly risk-averse, since capacity is set so high above the ‘perfect-foresight’ optimum. A limited knowledge of the risk informs the safety factor, leading to technologies that may never operate at full capacity or demand that might not be fully met; a designer cannot know beforehand whether either case will transpire. If additional cost is to be incurred in light of uncertainty, then the resilience of the resulting system should be well understood, to justify that increase.

Methods do exist for optimising energy systems to improve their resilience to uncertainty in system parameters. When implemented, these methods consider resilience from the perspective of robustness, which is one of the many ways in which resilience may be explored (Woods, 2015). Indeed, a key method, if the uncertain parameters are independent and well-bounded, is **robust optimisation (RO)**. Zugno *et al.* (2016) considered uncertainty in both heat demand and electricity market price in a day-ahead unit commitment RO model. Similar unit commitment studies have increased system robustness to weather, pricing and load uncertainty (Kazemzadeh *et al.*, 2017; Zhou *et al.*, 2018). However, both the strict requirement on uncertainty bounding and the computational burden of RO mean it is only viable for planning small district-scale energy systems (e.g. Akbari *et al.*, 2016; Majewski *et al.*, 2017).

Instead, uncertainty is commonly expressed by distinct scenarios which are then handled by **scenario optimisation (SO)**<sup>1</sup>. In SO, each scenario is optimised in parallel, such that investment decisions are robust to each scenario realisation. If the scenarios describe the uncertainty space sufficiently well, then the investment decision will be robust beyond the realisation of the scenarios for which they are optimised.

Generating scenarios for SO depends on both data availability and a modeller’s desire to fully describe the uncertainty. A better description of uncertainty will inevitably result in a more robust system. Whether encapsulating the uncertainty in weather, pricing, technology characteristics, or demand, three principal methods are used: physics-led, data-driven, or assuming some level of variance from the average. Of course, the latter is the easiest to incorporate. Often it is applied to a uniform distribution, where the choice of variance varies from  $\pm 10\%$  (Kuznetsova *et al.*, 2014; Pazouki and Haghifam, 2016; Pazouki *et al.*, 2014) to  $\pm 30\%$  (Ahn and Han, 2018) or  $\pm 40\%$  (Yokoyama and Ito, 2002). Other parametric distributions are also used when assuming variance, particularly Gaussian (Fuentes-Cortés *et al.*, 2016; Vahid-Pakdel *et al.*, 2017; Yang *et al.*, 2017), but also Weibull (Vahid-Pakdel *et al.*, 2017) and Kumaraswamy (Narayan and Ponnambalam, 2017). To generate scenarios, Monte Carlo simulations stochastically sample from within the assumed distributions. Understandably, the choice of variance significantly influences optimisation results (Ahn and Han, 2018; Yokoyama and Ito, 2002) and thus cannot be applied without sufficient validation.

Incorporating uncertainty from bottom-up, physics-led (engineering) models requires a detailed understanding of the physical laws to which a parameter must adhere. For instance, the thermodynamic interaction of building fabric, external conditions, and occupancy must

---

<sup>1</sup>‘Scenario’ and ‘stochastic’ are often used interchangeably to describe the same optimisation procedure. As stochastic optimisation may refer to other methods of handling stochastic variables, the distinction is made clear in this thesis by describing the method as scenario optimisation.

be modelled to understand thermal demand, while meteorological analysis is required to model weather. Stochastically varying input parameters prior to each simulation will create a set of future scenarios. In SO, only demand scenarios have been hitherto generated using engineering models. To do so, parameters such as material properties, ventilation rates, appliance energy consumption, and building occupancy are varied stochastically (Mavromatidis *et al.*, 2018b; Rezvan *et al.*, 2013). In such cases, the individual parameters are assumed to vary according to a normal distribution with assumed variance, which may cause validation concerns similar to applying the variance directly to demand. Indeed, specifying the uncertainty surrounding each engineering model parameter can be onerous (Zhao and Magoulès, 2012), especially when a number of different building types are included in a district (Reinhart and Cerezo Davila, 2016). Simplifications of this approach consider a smaller set of parameters to infer building-level energy demand, such as appliance usage from time-use surveys (Baetens and Saelens, 2016) or typical consumption patterns (Sharafi and ElMekkawy, 2015), and correlation with heating degree days (Sharafi and ElMekkawy, 2015).

In contrast, top-down data-driven models use historical data to identify inherent variability. They implicitly consider the complex interaction of all influencing parameters (predictors). If a sufficient amount of data is available, it is possible to understand the impact of each predictor by casting the data into statistical models (Braun *et al.*, 2014; Hong and Jiang, 1995; Jain *et al.*, 2014; Zhao and Magoulès, 2012). Sufficient data is often available for historical weather, leading to globally validated statistical models of wind and solar power availability (Pfenninger and Staffell, 2016; Staffell and Pfenninger, 2016). If there is insufficient data, the reliability of statistical models to represent the relationships between predictors and optimisation parameters can be limited. Alternatively, it is possible to sample historical data directly. In such cases, the emphasis is less on what *causes* a parameter to vary and more in capturing the variation such that it need not be assumed. With measured data from an office building, Gamou *et al.* (2002) stated that energy demand in any given hour can be described by a normal distribution, where there is a 95% probability of samples remaining within  $\pm 20\%$  of the mean. More recent studies have used this exact methodology to generate demand scenarios for optimising energy systems of hospitals (Yang *et al.*, 2017; Zhou *et al.*, 2013). Similarly, using six years of wind power data, Bruninx and Delarue (2016) showed that forecasting errors fitted better to the Lévy  $\alpha$ -stable distribution than previously assumed normal, Weibull or  $\beta$  distributions. At its most simple, historical data can be used directly; Bucciarelli *et al.* (2018) used 100 recorded days of energy demand as parallel scenarios for system optimisation.

A common shortcoming across current data-driven approaches for scenario generation is that autocorrelation is not considered. Energy demand at any particular hour is often influenced by the demand in other hours of the day, while large meteorological weather systems lead to spatial interdependence of renewable energy generation (Grams *et al.*, 2017). The uncertainty described by scenarios without considering autocorrelation can thus be misleading, and result in incorrect system design. Spatial variations in renewable energy provision is unlikely to impact district-scale systems. Indeed, no existing studies consider

each node to have distinct weather patterns. However, temporal autocorrelation of demand is particularly pertinent. Mavromatidis *et al.* (2018b) identified this as a key issue, which was solved in their study by use of a bottom-up engineering model for generating demand scenarios. However, scalability of this approach is limited.

Furthermore, the system design resulting from the consideration of uncertainty is rarely validated on its ability to mitigate risk. If a change in investment is required, which may inevitably increase system costs, the benefit of that system should be quantified. Therefore, it is important to test systems against unseen data, so called ‘out-of-sample’ scenarios, which represent possible realisations of the uncertain parameters under consideration (Conejo *et al.*, 2010). Akbari *et al.* (2016) generated ten out-of-sample scenarios to test the investment decisions resulting from a deterministic and a RO model, finding that the RO result reduced the standard deviation of unmet demand across the scenarios. Such a result implies an increase in system robustness, since greater invariance to parameter perturbations has been achieved (Alderson and Doyle, 2010). Validating with ‘out-of-sample’ tests is not common, however, which exposes a clear shortcoming of many recent studies (Ahn and Han, 2018; Majewski *et al.*, 2017; Mavromatidis *et al.*, 2018b)

In this chapter, a new three-step method to handle demand uncertainty in district-scale energy optimisation is presented, followed by its extension to mitigate the risk of power interruptions. The proposed method is data-driven and hence, unlike detailed bottom-up engineering models, scalable for application to large districts. The shortcoming of current data-driven models is overcome by sampling multivariate nonparametric representations of historical demand. Therefore, it is possible to account for temporal autocorrelation *and* skewness of demand data around the mean value at any given hour.

In the following section, the principal steps of the proposed method are detailed: (1) Data-driven scenario generation, (2) Scenario reduction, and (3) Scenario optimisation. Two case studies are used to demonstrate the application of this method: case studies 3 and 4 of this thesis. They are both introduced in detail in Section 5.2. Case study 3 is an illustrative district of 17 buildings in Bangalore, India, in which there is a proposed district cooling network which would be powered by a district combined cooling, heat and power plant (CCHP) or large-scale energy centre ECh (EC-ECh). Case study 4 is based on a real district of 46 buildings proposed for development in Cambridge, UK, in which there is a proposed district heating system which would be powered by a combined heat and power plant (CHP) or large-scale ground source heat pump (GSHP). In both case studies, scenario generation is undertaken using historical energy demand collected from existing buildings of the same archetype and the same geographic location.

The result of applying the three-step method to case studies 3 and 4 is given in Section 5.3. The degree to which robustness has been incorporated in the optimised designs is validated by the use of ‘out-of-sample’ tests in Section 5.4. Finally, an extension to the proposed method is analysed in Section 5.5, in which the impact of power interruptions in case study 3 is assessed and measures to reduce this risk are explored.

## 5.1 Handling uncertain demand

### 5.1.1 Scenario Generation

Particularly at a masterplanning level, little may be known about buildings within a district, other than their intended use and floor area. Using high-resolution historical data of other buildings representing similar use, a multidimensional search space can be created to describe the possible demand profiles for the district. To create a search space from the available data, terminology can be borrowed from machine learning by considering ‘features’ and ‘observations’. In the context of energy demand, features are the individual measurable properties that are being observed. Observations are the existing data describing distinct instances of those features. Features are the consumption values in each hour of a day, which have been observed for all historical days for which there is data. Over one year, there would be 24 features and 365 observations.

Clustering of observations into independent search spaces can help ensure that samples are more realistic. For example, weekend and weekday electricity consumption will differ in an office space. Seasons, academic term times, and months are all subjective clusters that could be chosen. There may be other, unknown metrics by which observations can be clustered programmatically, such as by K-means and hierarchical clustering (Pfenninger, 2017a). The advantage of subjective clustering is the ability to map those clusters into future years, which cannot readily be done with clustering algorithms. Typical days are commonly used in [mixed integer linear programming \(MILP\)](#), reducing the length of the time dimension from its full scale (e.g. one year - 8760 hours) to anything from three (Mehleri *et al.*, 2012) to six (Omu *et al.*, 2013) or twelve days (Ameri and Besharati, 2016). Clustered observations form search spaces which represent ‘typical’ days in the year.

The shape of a search space depends not only on the set (or subset) of observations, but also the method by which interpolation is undertaken to create a continuous surface from discrete data. Previous studies have considered well-conditioned demand, which can be described by a parametric distribution, such as a multivariate normal curve (Sun, 2014; Ward *et al.*, 2016b). However, daily demand profile sets generally do not fit a perfect Gaussian profile. This is especially the case when observations are acquired from multiple buildings within a single archetype. There can be sub-clusters of demand profiles with various local peaks in the distribution. Indeed, an initial analysis shows that demand is not well conditioned; i.e. it cannot be well described by a parametric distribution. From Figure 5.1 it can be seen that clear clusters of profiles are lost when sampling parametrically, assuming multivariate normal distributions. The symmetrical nature of multivariate normal sampling has also led to areas of high profile density to become the sample mean as well as the mode.

Nonparametric sampling can be more representative of the demand. However, as no assumptions are made about the shape of the input data, the search spaces resulting from nonparametric sampling are heavily dependent on the input data. One method to overcome this is to have large training and validation sets to tune the relevant hyperparameters. Two

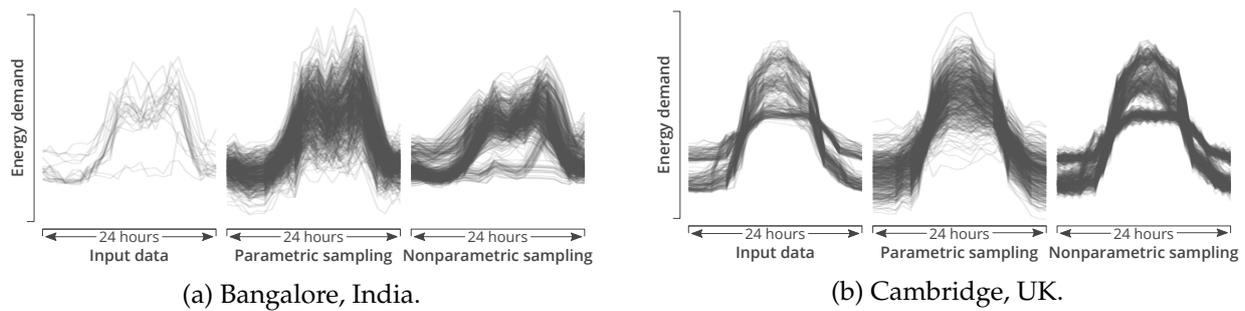


Figure 5.1 Comparison of parametric and nonparametric sampling methods as a means to produce daily profiles for demand. Profiles are at an hourly resolution. Input data given here is for one typical day, from which 500 profiles were drawn using each of the sampling methods.

approaches are presently discussed in which a smaller data set may still be valid: multi-building data sets and [functional principal component analysis \(fPCA\)](#).

Replication of the demand profiles of one building is not desired when sampling a search space. To avoid this, data describing multiple buildings within an archetype can be sampled. Then, when sampling for an individual building, stochastic profiles will be unlikely to duplicate any one of the input profile sets. When little is known about a building other than its archetype, there is validity in this approach. The demand in an archetypal building could replicate the profiles in any of the input data buildings in that archetype. In fact, combining data on multiple buildings was used by the masterplanners when assigning archetypal demand to the Cambridge case study district (Cambridge City Council, [2016](#)).

As data from multiple buildings inform the search space for a single archetype, when two consecutive days are sampled from the profiles of different buildings in the same archetype, wildly varying demand may be observed from one day to the next. Much like the intraday temporal autocorrelation of demand, interday autocorrelation is also necessary. To account for this, before stochastically sampling profiles, energy intensity for each input building can be normalised by the maximum demand recorded for that building. After sampling, modelled buildings had their demand scaled by a randomly assigned normalisation factor from all those available for their archetype.

Without multiple buildings as input data, another approach must be taken to ensure samples are representative of the archetype and not only the input building. This can be done in [fPCA](#) since demand for a typical day is treated as a function of time, as opposed to being discrete data points (Ramsay and Silverman, [2005](#)). It provides a mathematical definition of the shape of the curve in terms of a number of functional Principal Components (PCs) which are the same for all the data samples and describe particular features of the data. A set of weightings, or ‘scores’, is associated with each PC. The scores describe mathematically the contribution of each PC to the overall day’s demand profile, per end-use. A search space is then created for each typical day by using the set of scores associated with them (Lu *et al.*, [2016](#); Ward *et al.*, [2018](#); Ward *et al.*, [2016a](#)). Thus, generated demand profiles retain the primary *sense* of the input data without it being exactly reproduced.

In both methods of preparing the input data, [kernel density estimation \(KDE\)](#) can be used to create a [probability density function \(PDF\)](#). In KDE, a kernel (e.g. normal distribution) is

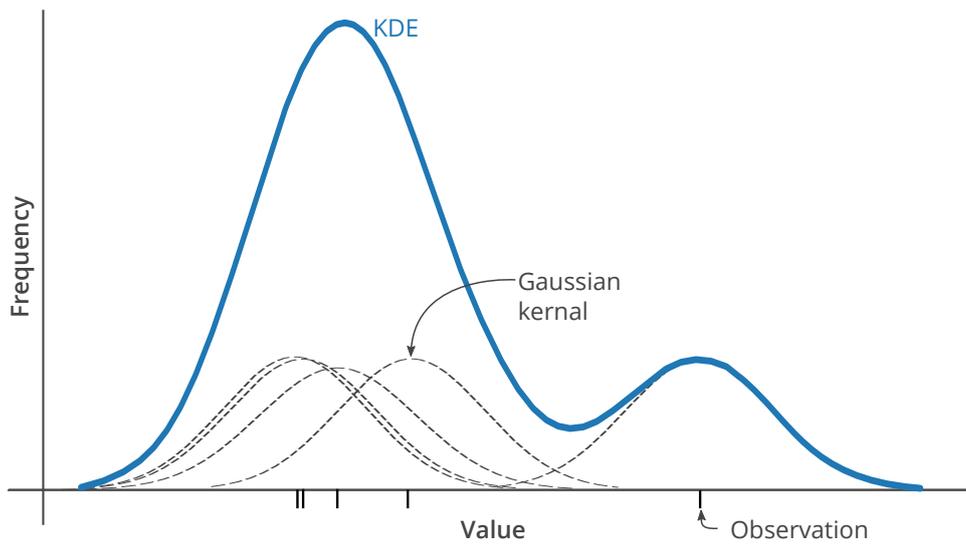


Figure 5.2 Example of KDE, applied to a single feature observation set. A gaussian kernel has been applied to each observation in this case.

applied to each feature in the observations, and the overlay of all these individual kernels represents the full data set. Figure 5.2 represents this with one feature<sup>2</sup>, which is a nonparametric distribution due to its lack of a single peak. There are two important hyperparameters which dictate the efficacy of KDE: the kernel and the bandwidth. The kernel is the shape of the density function applied to each observation when constructing the full PDF. It is standard to use a Gaussian kernel (as used in Figure 5.2), but ‘top hat’, ‘triangular’, and ‘Epanechnikov’ are among other kernel choices (Turlach, 1993). The bandwidth is the scale of smoothness applied to each kernel, akin to the standard deviation of a normal distribution. If bandwidth = 0, the resulting PDF will have non-zero values only at points corresponding to the input data. As the bandwidth increases to infinity, the PDF converges on a uniform distribution, with infinite variance. Sampling from either of these extreme cases is not advisable. Instead, it is desirable to find the lowest possible bandwidth that fits a training data set and can reproduce an independent validation data set on sampling. If there are too few observations for training and validation,  $k$ -fold cross-validation can be used in bandwidth and kernel selection. In  $k$ -fold cross-validation, the full data set is randomly partitioned into  $k$  subsets. One subset is retained for validation while the remaining  $k - 1$  sets are used for training. The process is repeated  $k$  times, such that all subsets are used for validation and training (Refaeilzadeh *et al.*, 2009).

In Figure 5.1 the impact of multivariate nonparametric sampling can be seen using multi-building sets (Cambridge) and fPCA (Bangalore), both sampled from a KDE-generated PDF. fPCA samples were generated by Ward *et al.* (2018) while multi-building sets were prepared as part of this research. The input data is better represented by nonparametric, rather than parametric sampling; The demand profiles cover a wider space beyond the range of the input data, while sub-clusters across the data are still apparent.

<sup>2</sup>the KDE of 24 features, for each hour in the day, would form a 24-dimensional PDF, which is impossible to visualise.

In this chapter, [fPCA](#) has been used to generate scenarios for the Bangalore study and multi-building sets for the Cambridge study. 500 demand profiles are generated per typical day, per end-use, and per building archetype in kWh/m<sup>2</sup>. Annual hourly demand scenarios per building are sampled from these stochastic profiles for a reference year (Cambridge: 2015, Bangalore: 2016), such that no profile is duplicated between days in the year, or between buildings in the district. A summary of the 500 scenarios is given in [Figure 5.8](#) when discussing the results of applying this method.

### 5.1.2 Scenario reduction

A large number of probabilistic demand scenarios per building can result in intractability of a district energy system optimisation model. Accordingly, the selection of the ‘right’ subset of scenarios becomes an important step, especially in scenario optimisation models. This subset of scenarios should be representative of the variations across the larger scenario set without requiring the whole set to be included in the optimisation model. [Conejo \*et al.\* \(2010\)](#) proposed the use of the fast-forward algorithm to reduce the number of scenarios. Two primary variants of the method are proposed, both of which aim to reduce the *Kantorovich distance*<sup>3</sup>. Both variants apply a cost metric to each scenario  $s$  in the scenario set  $S$ , from which a subset  $S'$  is chosen based on the minimisation of the difference in the probability distributions describing the costs in  $S$  and  $S'$ .

The two variants of scenario reduction proposed by [Conejo \*et al.\* \(2010\)](#) differ on the cost metric applied to each scenario. In the first, a key metric describing the scenarios is selected. This might be the maximum hourly demand per scenario or the total demand over the entire year. The metric is selected subjectively as a measure that has the biggest impact on the objective function. The first variant is considered less computationally intensive to apply and has been used for scenario reduction in existing [SO](#) studies ([Good and Mancarella, 2017](#); [Mavromatidis \*et al.\*, 2018a](#)). The second variant requires that a non-probabilistic optimisation model is run for each scenario independently. These independent models are a formulation of the [SO](#) model which do not consider uncertainty. They are relatively fast to solve and can be run in parallel on a high-performance cluster in a matter of minutes. The objective function value calculated for each scenario is then used in the Kantorovich distance calculation.

A refinement of the second variant, proposed by [Bruninx and Delarue \(2016\)](#), is used in this study. [Conejo \*et al.\* \(2010\)](#) only considered optimisation of operation costs, fixing the investment cost as a parameter in each of the independent models. [Bruninx and Delarue \(2016\)](#) included investment and operation costs as decision variables. All decision variables are therefore part of the independent model optimisation, but binary and integer constraints were not included for computational efficiency. The independent models in this study retain the binary ‘purchase’ constraints applied to investment decisions. Because typical days are used, as against the full time series of annual demand, each independent optimisation runs within a reasonable solution time ( $O(100s)$  on a high-performance cluster).

The process for scenario reduction can be thus summarised as follows:

<sup>3</sup>For a detailed mathematical formulation, readers are referred to ([Conejo \*et al.\*, 2010](#); [Römisch, 2009](#)).

1. Optimise the objective function for each scenario in parallel, minimising system cost (investment and operation) for each case independently.
2. Select a subset of scenarios to represent the 500 input scenarios, by minimising the Kantorovich distance between the probability density of their objective function values to that of the full scenario set.
3. Assign each scenario in the full set to the closest (by probability distance) of the scenarios in the reduced subset, weighting each reduced scenario by the number of scenarios it represents.

The result of this process is a scenario subset that can be used for tractable scenario optimisation. It is applied in the same manner for both the Bangalore and Cambridge case studies. The selected reduced scenarios are detailed alongside results for [scenario generation \(SG\)](#) and scenario optimisation, in section [5.3](#).

### 5.1.3 Scenario optimisation

Once reduced scenarios are derived, the uncertainty described by these scenarios can be dealt with by [SO](#). A [SO](#) model has two stages. The first involves finding the optimal technologies and their capacities, irrespective of their ability to meet variability in demand. In the second stage, the optimal technologies are reassessed for their ability to meet variability in energy demand across all reduced scenarios. The impact of a single scenario is weighted by its probability of occurrence,  $W_{s^t}$  (Equation [5.1](#)), such that low probability scenarios may operate in a high cost manner without incurring a large penalty on the overall objective function. In fact, the strict requirement that demand must be met can be relaxed in favour of a financial penalty associated with unmet demand. The slack variables in the objective function act to allow unmet demand or excess supply, but they incur a high cost,  $bigM$ . Having unmet demand is a risk, but if it occurs in a low probability scenario then its impact on the objective function may be small.

$$\begin{aligned}
min \quad & TSF \times \sum_{n,x,y} \left( \mathbf{y}_{n,x} \times cost_{n,x,k}^y \right) \\
& + \sum_{n,x,c,s^t,t} W_{s^t} \times \left( \mathbf{E}_{n,x,c,s^t,t}^+ \times cost_{n,x,k,t}^{prod} - \mathbf{E}_{n,x,c,s^t,t}^- \times cost_{n,x,k,t}^{con} + \mathbf{E}_{n,x,c,s^t,t}^{ex} \times cost_{n,x,k,t}^{ex} \right) \quad (5.1) \\
& + bigM \sum_{n,c,s^t,t} W_{s^t} \times \left( \mathbf{slack}_{n,c,s^t,t}^+ - \mathbf{slack}_{n,c,s^t,t}^- \right)
\end{aligned}$$

Where  $n$  = node,  $x$  = technology,  $c$  = energy carrier,  $k$  = monetary cost class,  $t$  = timestep,  $y$  = capacity type ( $\hat{E}$ ,  $\hat{S}$ , etc.),  $s^t$  = scenario, and  $TSF$  = the timeseries scaling factor. Refer to page [42](#) for more information on model decision variables and sets.

[SO](#) can be approached with varying levels of risk aversion. Here, risk is the impact of the worst case realisations of the future, i.e. those which incur a high cost. If a modeller is 'risk-neutral', Equation [5.1](#) may be used: scenario impact is proportional to its probability

of occurrence, irrespective of whether it is high or low risk. Instead, a modeller may choose to be ‘risk-averse’. A risk measure must then be applied to the objective function, to disproportionately weight scenarios which will lead to a relatively high system cost. One such risk measure is **conditional value at risk (CVaR)**, which allows particularly unfavourable scenarios to be severely penalised within the objective function. There are as many possible system costs in a **SO** model as there are scenarios, all of which describe a **PDF**. The **CVaR** describes the sum of the expected cost above a given confidence level  $\alpha \in [0, 1)$  in this **PDF** (Figure 5.3). As it concentrates on the right-hand tail of the distribution, it is a risk measure that is only influenced by the worst-case scenarios. It is an extension of the **value at risk (VaR)** measure, which minimizes the cost at the confidence level  $\alpha$ , that is better suited to linear optimisation models (Rockafellar and Uryasev, 2002).

Maurovich-Horvat *et al.* (2016) used risk-averse **SO** to account for uncertain spot pricing in the optimal operation of a **CHP**. **CVaR** is used to penalise the use of the **CHP** when the electricity export price is low. Similarly, Bukhsh *et al.* (2016) considered risk-averse wind power operators in energy market clearing. The need to curtail wind power or to operate back-up generators, due in both cases to incorrect prediction of wind power availability, is financially penalised within the **CVaR** component of the objective function.

The risk-averse objective function given in Equation 5.2 adds **CVaR** to Equation 5.1. **CVaR** of a scenario is estimated using  $\eta \in [0, \infty)$ , weighted by the probability of occurrence of that scenario,  $W_{s'}$ , and calculated across scenarios by combination with the model-wide **VaR**,  $\xi$ . Model-wide parameters  $\alpha$  and  $\beta$  describe the confidence interval and risk aversion hyperparameters, respectively. If  $\beta = 0$ , Equation 5.2 becomes Equation 5.1.  $\xi$  and  $\eta$  are connected to existing decision variables in the model by the constraint given in Equation 5.3. The optimisation looks to balance  $\xi$  and  $\eta$  to minimise the risk-aversion penalty. If  $\xi$  is low, the impact of risk-aversion is low but  $\eta$  is necessarily higher.  $\eta$  is only ever non-zero for scenarios with a penalty cost greater than  $\xi$ , but it has a  $N$ -fold impact on the penalty, where  $N$  is  $\frac{1}{1-\alpha}$ . If  $\alpha$  is 0.9,  $N$  becomes 10. Therefore, keeping even the worst case cost low is beneficial as it allows  $\xi$  to remain low without requiring  $\eta$  to ever be non-zero.

$$\begin{aligned}
\min \quad & TSF \times \sum_{n,x,y} \left( \mathbf{y}_{n,x} \times \text{cost}_{n,x,k}^y \right) \\
& + \sum_{n,x,c,s',t} W_{s'} \times \left( \mathbf{E}_{n,x,c,s',t}^+ \times \text{cost}_{n,x,k,t}^{prod} - \mathbf{E}_{n,x,c,s',t}^- \times \text{cost}_{n,x,k,t}^{con} + \mathbf{E}_{n,x,c,s',t}^{ex} \times \text{cost}_{n,x,k,t}^{ex} \right) \\
& + \text{bigM} \sum_{n,c,s',t} W_{s'} \times \left( \mathbf{slack}_{n,c,s',t}^+ - \mathbf{slack}_{n,c,s',t}^- \right) \\
& + \beta \left( \xi + \frac{1}{1-\alpha} \sum_{s'} (W_{s'} \eta_{s'}) \right)
\end{aligned} \tag{5.2}$$

Where  $n$  = location,  $x$  = technology,  $c$  = energy carrier,  $k$  = monetary cost class,  $t$  = timestep,  $y$  = capacity type ( $\hat{E}$ ,  $\hat{S}$ , etc.),  $s'$  = scenario, and  $TSF$  = the timeseries scaling factor.

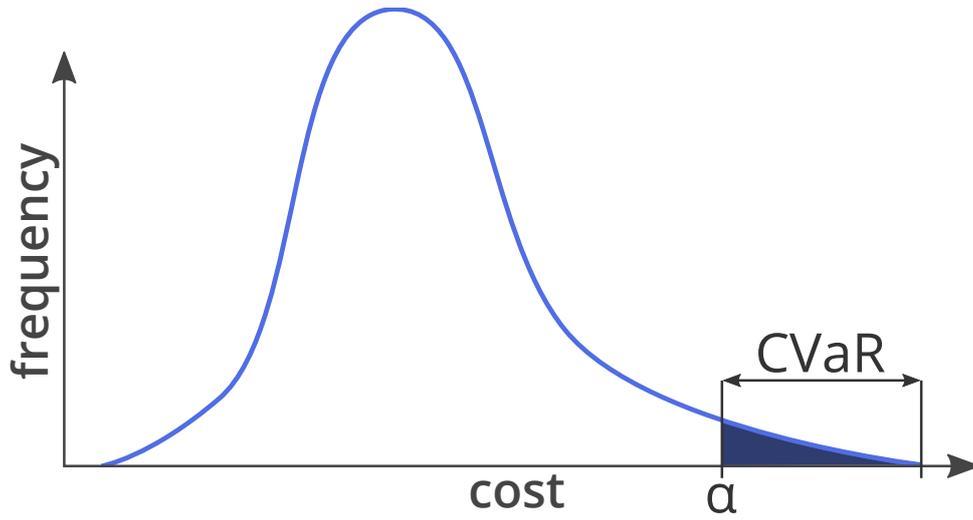


Figure 5.3 Representation of the CVaR risk measure, which is the expected value of observations above a given percentile  $\alpha$  in the PDF describing all observations in a dataset.

$$\sum_{n,x,c,s',t} W_{s'} \left( \mathbf{E}_{n,x,t}^+ \times \text{cost}_{n,x,k,t}^{\text{prod}} - \mathbf{E}_{n,x,c,s',t}^- \times \text{cost}_{n,x,k,t}^{\text{con}} + \mathbf{E}_{n,x,c,s',t}^{\text{ex}} \times \text{cost}_{n,x,k,t}^{\text{ex}} \right) + \text{bigM} \sum_{n,c,s',t} W_{s'} \times \left( \text{slack}_{n,c,s',t}^+ - \text{slack}_{n,c,s',t}^- \right) - \zeta \leq \eta_{s'} \quad (5.3)$$

Where  $n$  = location,  $x$  = technology,  $c$  = energy carrier,  $k$  = monetary cost class,  $t$  = timestep,  $y$  = capacity type ( $\hat{E}$ ,  $\hat{S}$ , etc.),  $s'$  = scenario, & TSF = the timeseries scaling factor.

In this chapter, both risk-neutral and risk-averse SO studies are undertaken. This thesis focusses on the use of MILP as the optimisation technique. However, other optimisation techniques can be applied within the proposed SO framework.

### Quantifying risk of unmet demand

SO studies which aim to minimise risk may or may not introduce a new penalty cost associated with risk. Particularly when there is no CVaR component considered, the objective function remains similar to the single scenario case, only with a weighted sum of the different scenarios (Bucciarelli *et al.*, 2018; Mavromatidis *et al.*, 2018b; Pazouki and Haghifam, 2016). Particularly when a risk component is added to the objective function, such as CVaR, some studies choose to allow scenarios to perform badly, if they are willing to pay the price. For instance, Bukhsh *et al.* (2016) used the costs of dispatchable, expensive diesel generators in cases of insufficient wind supply. Scenarios with low probability of occurrence may risk committing to selling more energy from wind power, as the impact on the objective function of requiring to use diesel generators will be low. Similarly, Maurovich-Horvat *et al.* (2016) used the higher cost spot electricity and gas prices as a penalty for incorrectly purchasing energy from the futures market.

Including a penalty term that would otherwise not be chosen in a single scenario model is useful. It allows models to better consider the probability of the scenarios in making the decision of how best to meet demand. In this chapter, the risk incurred in the system is that

demand will not be met. As such, the risk of unmet demand should be quantified. The energy balance in an energy system explicitly does not allow unmet demand. However, it does tend to include 'slack' variables, which impose a high cost on meeting demand, but allows for them to be used if, on realising an uncertain future, demand cannot be met. The use of diesel generators by Bukhsh *et al.* (2016) could be considered a slack variable, given that such technologies would not be chosen if no demand uncertainty existed.

Where a physical entity can be used as a slack variable, the associated (usually high) cost of using that entity is known. In the case of the models represented in this chapter, the penalty represents a virtual source of energy supply, to ensure all demand is met. This might translate in reality to a diesel generator for electricity, window-mounted air conditioning for cooling, or electric radiators for heat. None is likely to incur sufficiently high costs to limit their use to only extreme cases. Nor should they. If demand is not met by the installed technology portfolio then it means that these additional technologies have not yet been purchased. Given the unexpected nature of the unmet demand, making these purchases would inevitably be retroactive and may not be made at all, if there is no understanding of how often further instances of unmet demand are likely to occur. This cost therefore must monetize and include factors such as commercial losses and loss in productivity. Indeed, losses in commercial sales due to grid-based power outages can be up to 2% in Bangalore (Dollar *et al.*, 2005).

Deciding on the cost of unmet demand in this research is thus subjective. In single-scenario models, unmet demand is valued at  $1 \times 10^{10}$ , to ensure it only exists to provide model feasibility. For application as a penalty in a feasible SO model, the cost of unmet demand was tested using the Bangalore case study introduced in this chapter. The model was optimised several times, with the cost of unmet demand increasing in each run in increments of a factor of ten, from 100 INR/kWh to  $10^7$  INR/kWh. The result, as expected, is a decrease in unmet demand with increased cost.

Figure 5.4 shows the magnitude of decrease in unmet demand across 16 scenarios used in the test, for the penalty range between 100 INR/kWh and  $10^6$  INR/kWh. At  $10^7$  INR/kWh there is no unmet demand. Unmet demand is found uniformly across scenarios at penalty levels  $10^2$  and  $10^3$  INR/kWh. Only from  $10^4$  INR/kWh is there a particular weighting of unmet demand onto lower probability scenarios.

From  $10^5$  INR/kWh, the investment decisions were also seen to stabilise. Although applying a lower cost to unmet demand ensures that highly unlikely scenarios do not dictate the investment decision, at costs lower than  $10^5$  INR/kWh the reduction of technology capacity was found to adversely impact most of the scenarios. As such, an unmet demand penalty of  $10^5$  INR/kWh has been applied to the Bangalore case study and, following the exchange rate of  $\sim 100$  INR to GBP, an unmet demand penalty of  $10^3$  INR/kWh has been applied to the Cambridge case study. Both are orders of magnitude greater than the highest cost for purchasing electricity, but in particularly low probability scenarios it may prove cost effective to shed load to avoid unnecessarily increasing capacity investment.

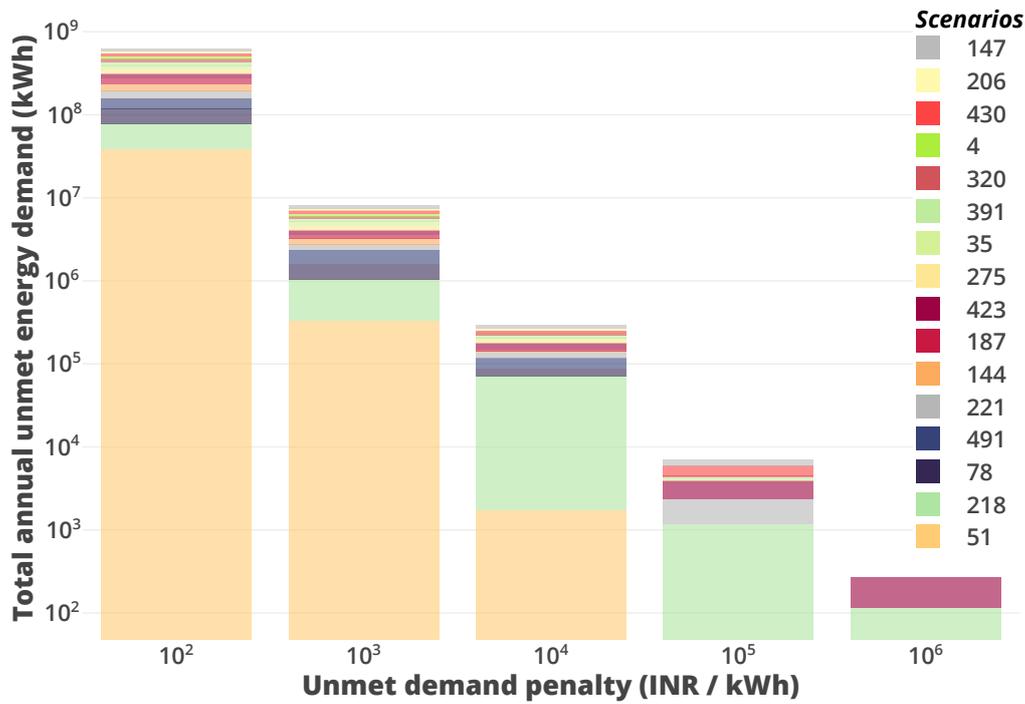


Figure 5.4 Comparison of quantity of unmet demand incurred in various scenarios for penalty rates (*BigM* in Equations 5.2 and 5.1) from 100 to  $10^6$  INR/kWh, applied to a test model of the Bangalore case study district. **NOTE:** Scenarios are given in ascending order of probability, but as the y axis is a  $\log_{10}$  scale, lower probability scenarios appear to have a greater contribution than is actually the case.

## 5.2 Model configuration

### 5.2.1 Case studies

The two case studies represent very different techno-economic-geographic contexts and hence test different types of demand (cooling dominated versus heating dominated) and associated technologies. Their energy systems are optimised for minimising total costs (investment and operation, normalised to one year) following the same steps. However, a different manner of processing historical energy demand data for the scenario generation step is used for each case study. The reason for this is solely due to the difference in nature and volume of data available from the two sites. At the same time, it is true that historical data from buildings is never uniformly available across sites, and future applications of the proposed work may devise their own techniques for processing their demand data as long as they preserve key properties for estimating future possible demand scenarios. Piecewise linearisation of technology part-load consumption curves (detailed in Chapter 4) has not taken place in this chapter. Due to the greater complexity in the spatial dimension, and addition of the scenario dimension, only the use of rated efficiency proved tractable.

#### Bangalore, India

**District** A collection of office buildings within Bangalore, India have been selected, defining an illustrative district. Figure 5.5a shows each of these buildings, and the nodes used to represent them. Most nodes consist of several buildings, which are connected at the same point on the district cooling network. Building floor area (Table 5.1) has been inferred from

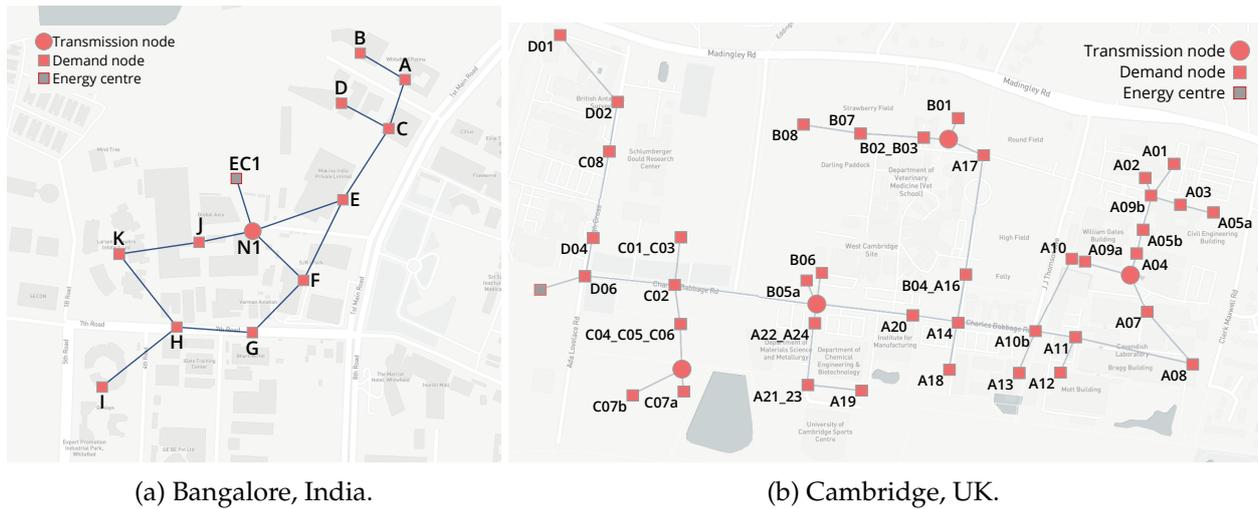


Figure 5.5 Case study districts.

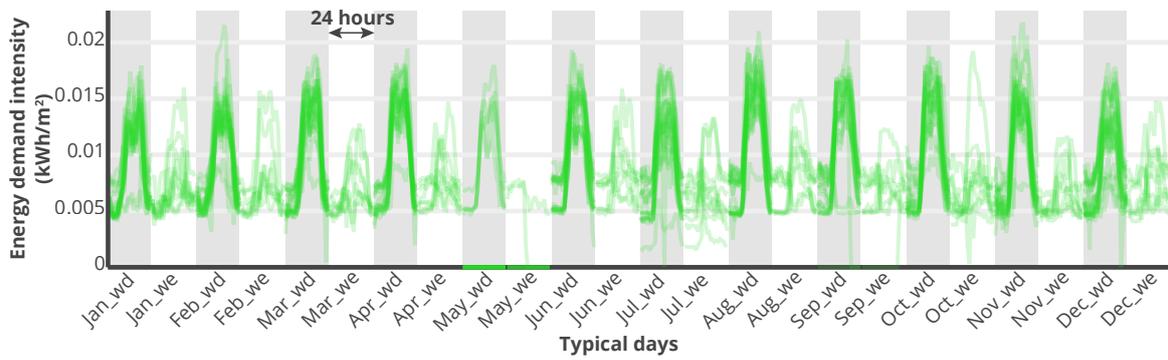
Table 5.1 Bangalore district node details. GIA = gross internal floor area. Areas given to three significant figures.

Node	A	B	C	D	E	F	G	H	I	J	K	EC1	N1
GIA ( $\times 10^3 \text{m}^2$ )	5.44	36.6	12.7	22.4	17.2	78.1	46.6	93.1	23.8	178	39.5	N/A	N/A
Roof area ( $\times 10^3 \text{m}^2$ )	2.72	6.10	3.16	5.60	8.59	11.16	11.65	18.61	5.96	22.31	7.90	N/A	N/A
Technologies	PV, ECh, DG, GridE											ECh, CCHP, StoreT, GridE	

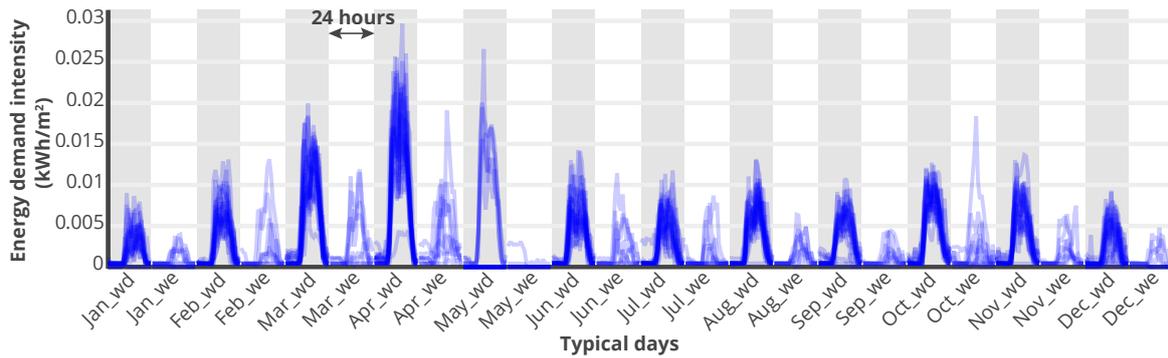
the external footprint and number of floors for each building. As the development is fictitious, no other information is known about these buildings; only their relative size and position is used to test the modelling approach.

**Demand** The demand data used in this study has been acquired from a single office building in Bangalore. Five-minute sub-metered data is available for this building for a range of end-uses. These end-uses include air conditioning, lighting, and uninterruptible power supply. The data was captured from December 2015 to November 2016 (inclusive) and is shown clustered by months and weekend/weekdays (sampling clusters) in Figure 5.6. Air conditioning electrical consumption is converted to cooling demand by using the variable refrigerant flow system **coefficient of performance (COP)** of 1.6. Current literature would suggest a **COP** between 3 and 4 (Lee *et al.*, 2018). The low **COP** is that which was recorded for the system in operation within the metered office building. Given that the system has been in place for many years, older literature suggests that it is not an unreasonably low **COP**. Xia *et al.* (2002) recorded a **COP** lower bound of 1.9, while Zmeureanu (2002) found their rooftop units to have operational **COP** of  $1.68 \pm 0.19$  and  $1.86 \pm 0.37$ , which compared particularly unfavourably to the rated **COP** of 2.9. To match the resolution of climate data, the five-minute metered consumption was resampled to hourly data.

There are clear trends visible in Figure 5.6 per sampling cluster, with weekends exhibiting a greater variation in demand. As expected, climate clearly affects cooling demand, with



(a) Energy type: electricity.



(b) Energy type: cooling.

Figure 5.6 Bangalore case study input demand profiles, grouped by sampling cluster and energy type. Profiles have low opacity such that darker sections indicate significant profile overlap. ‘\_we’ = weekend, ‘\_wd’ = weekday.

high April external temperatures causing high cooling requirements. In May, there are many days with zero demand. These probably correspond to building shutdown days in the summer holiday period. Some zero demand days may be caused by metering errors, but this top-down method for assessing demand data makes it difficult to assess whether they are erroneous points or truly zero-demand days. Stochastic profiles were generated per sampling cluster by Ward *et al.* (2018) using *fPCA*, the result of which is given in Section 5.3.1.

**Technologies** Table 5.1 shows the available technologies at each node. There is no requirement for any technology to be installed at any particular node, as the investment step of the optimisation will decide this. At a building level, *GridE*, *diesel generators (DGs)*, and *solar photovoltaic panels (PVs)* are possible technologies to meet electricity demand. In addition to the district cooling system, individual *electric chiller (ECh)* units can meet cooling demand. In the central energy centres, a large scale *ECh* or a *CCHP* can be installed. The *CCHP* is either fuelled by diesel (*D-CCHP*) or biomass (*B-CCHP*), whose waste exhaust heat is redirected through a *heat recovery absorption refrigerator* to produce cooling. *Thermal energy storage (StoreT)* is possible at the energy centre, but due to the relatively low energy density of cold water, *StoreT* is not considered at a building level. Technology costs are collated from various sources. Where costs specific to India were not available, values from the (UK specific) SPON’S mechanical and electrical services price book (AE-

Table 5.2 Cambridge district node details. GIA = gross internal floor area.

Node	desk_research (10 buildings)	lab_research (17 buildings)	desk_commercial (18 buildings)	lab_commercial (1 building)	Energy centre
GIA	86,774	176,173	181,908	9,473	N/A
Roof area	26,715	72,648	69,824	5,535	
Technologies	PV, NGB, ST, StoreE, StoreT			CHP, StoreT, GridE	

COM, 2015) have been used, assuming a currency conversion factor of 90:1 INR:GBP. More detail on the district and technology definitions is available in Section B.3 and online at <https://github.com/brynpickering/bangalore-calliope>.

### Cambridge, UK

**District** Unlike the Bangalore ‘representative’ district, the Cambridge district is based on intended development by the University of Cambridge. The West Cambridge site is a campus of the University, in which there already exists a number of academic, residential, leisure, and commercial buildings. The plan<sup>4</sup> is to construct 383,000m<sup>2</sup> of new floorspace, through a combination of greenfield and brownfield development (the latter directly replacing current buildings). According to the masterplan (Cambridge City Council, 2016), the district will have a 42GWh annual heating load, 70% met by a CHP, and 88GWh annual electricity load, 29GWh of which will be met by the same CHP. To determine this expected load, the buildings on the proposed site were categorised by archetypes, including ‘desk-based research’, ‘medium intensity laboratories’, and ‘high intensity laboratories’. Based on the archetype approach, and following consultation with Aecom, the contracted consultants for the energy plan of the West Cambridge site, four building archetypes are considered in this study: ‘Desk-based Commercial’, ‘Desk-based Research’, ‘Lab-based Commercial’, and ‘Lab-based Research’.

**Demand** The masterplanners used expert judgement and the mean annual demand of existing buildings to inform the expected demand of the proposed buildings within each archetype on the West Cambridge site. Similarly, buildings on the University estate have informed the expected demand within this case study district. Gas and electricity consumption data for 17 buildings have been accessed from across the University. An 80% efficient boiler is assumed available to meet heat demand from the incoming gas. As with the Bangalore case study, demand input is grouped by subjective sampling clusters. Many of the buildings are academic, so clustering of the data is based on weekends/weekdays and the University of Cambridge term dates (Table 5.3), leading to 12 sampling clusters. Different profile sub-clusters originating from different buildings within the same archetype are evident in Figure 5.7. In some cases, this can lead to an order of magnitude variation in the possible peak on a given day. Electricity demand profile shapes are more pronounced than heating demand profiles, although there is a pronounced morning heating peak in desk-based research/commercial buildings. On weekends, demand is lower but the profile shape is more sporadic, particularly in the ‘lab’ archetypes. This is probably caused by lab occupants

<sup>4</sup>More detail on the West Cambridge plan can be found at <http://www.westcambridge.co.uk/>

Table 5.3 Dates corresponding to Cambridge term times (Term) and vacations (Vac), as used to define sampling clusters for the Cambridge case study. Dates given for 2015 in *month-day* format. ‘Vac3’ dates wrap from December to January.

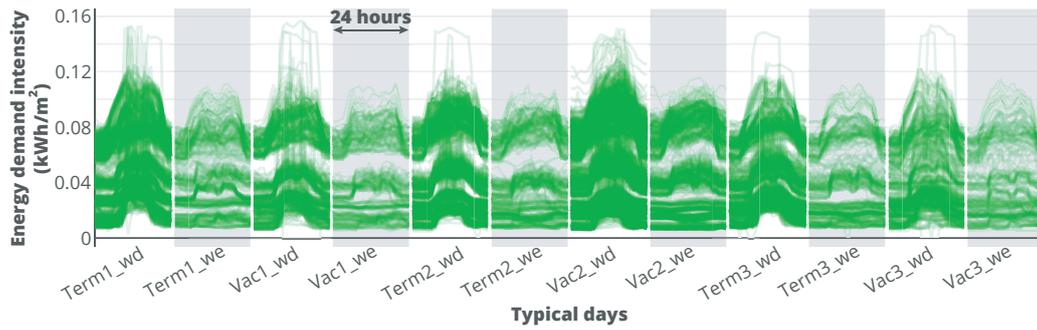
Term			Vac		
1	2	3	1	2	3
01-13 to 03-13	04-21 to 06-12	10-06 to 12-04	03-14 to 04-20	06-13 to 10-05	12-05 to 01-12

choosing to work on weekends and unsupervised, energy intensive lab experiments taking place over weekends. The buildings require less heating in summer (‘Vac2’), with hot water requirements being the likely cause of the remaining heat demand. Stochastic profiles were generated per sampling cluster and archetype using KDE, the result of which is given in Section 5.3.1.

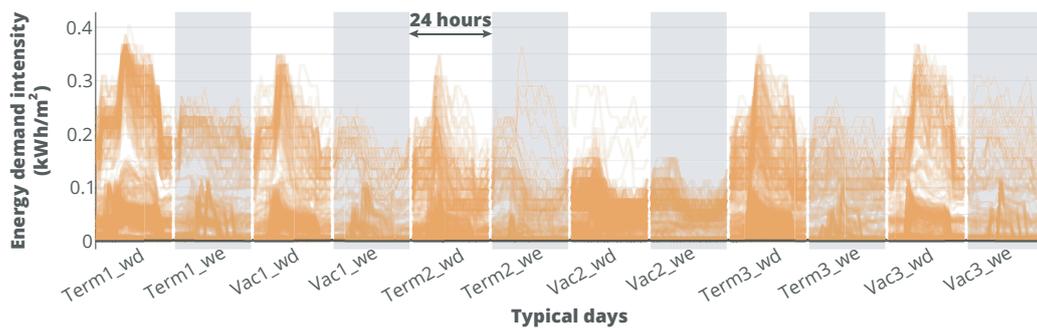
**Technologies** Buildings within the chosen archetypes are scattered across the development site. Figure 5.5b shows the proposed CHP would be based in an energy centre at the western edge of the district. The heating network follows the road network in connecting to the buildings. To account for the installation of gas pipework, a gas network cost is applied along the same route. To make the most of a possible energy centre, there is the possibility of a large-scale GSHP and StoreT to enter into the district system. Nodes in the district correspond to buildings of different archetypes, with roof area available for solar technologies and building-level technologies made available, if a district system is not favourable (see Table 5.2). Building-level heat demand can be met by natural gas boilers (NGBs) or solar thermal panels (STs), and can be stored using building-level StoreT. PV, GridE, and CHP output can meet electricity demand. Electricity can also be stored using electrical battery storage (StoreE) at a building level. Technology costs are primarily taken from the SPON’S mechanical and electrical services price book (AECOM, 2015). More detail on the district and technology definitions is available in Section B.4 and online at <https://github.com/brynpickering/cambridge-calliope>.

## 5.2.2 Case study uses

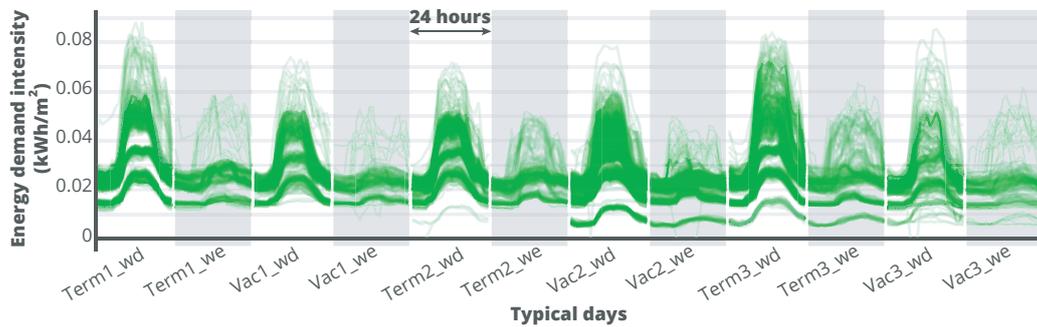
The objective functions in Equations 5.1 and 5.2 are subject to various constraints typical to energy system models, all of which are formulated within Calliope. Section 3.2 details the parameters, constraints, and decision variables used in all models in this thesis. Tracking storage between sampling clusters and SO was introduced to the Calliope framework for the studies in this chapter. Models were run on a high performance computing cluster, with optimisation undertaken by the CPLEX solver (v12.6.2) (IBM Corp., 2016).



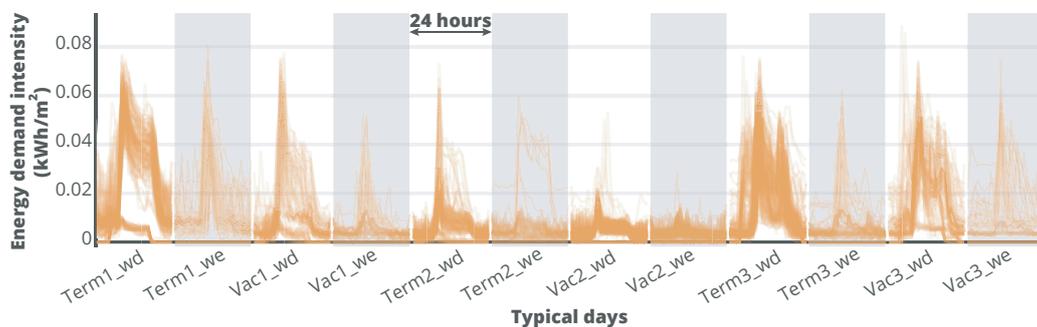
(a) Archetype: lab\_research; energy type: electricity.



(b) Archetype: lab\_research; energy type: gas.

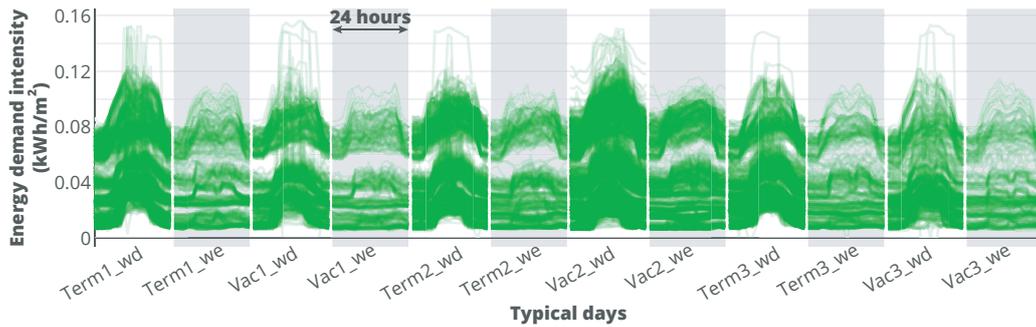


(c) Archetype: desk\_research; energy type: electricity.

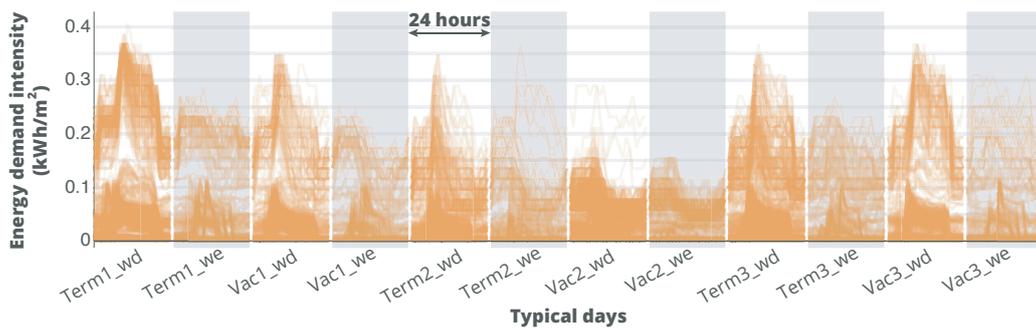


(d) Archetype: desk\_research; energy type: gas.

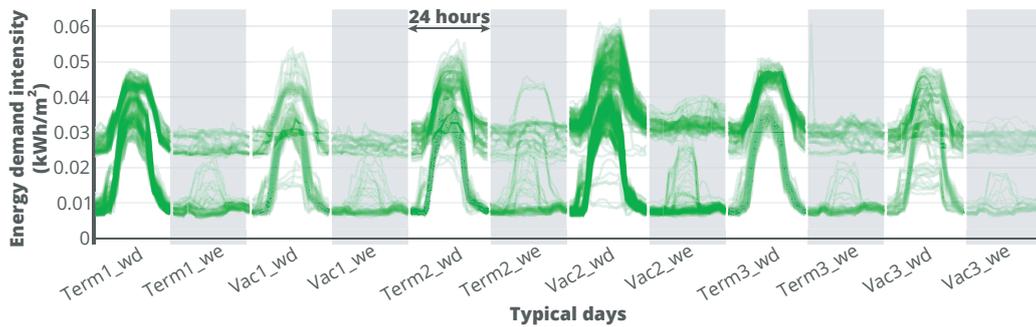
Figure 5.7 Cambridge case study input demand profiles, grouped by sampling cluster, archetype, and energy type. Profiles have low opacity such that darker sections indicate significant profile overlap. See Table 5.3 for sampling cluster dates and Table 5.2 for buildings per archetype. ‘\_we’ = weekend, ‘\_wd’ = weekday.



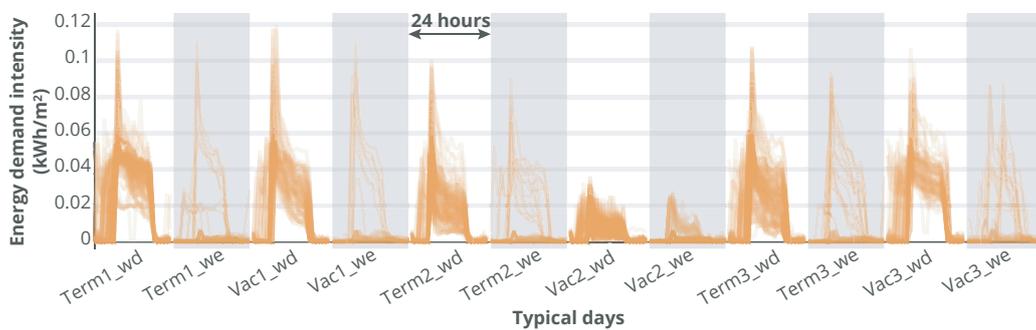
(e) Archetype: lab\_commercial; energy type: electricity.



(f) Archetype: lab\_commercial; energy type: gas.



(g) Archetype: desk\_commercial; energy type: electricity.



(h) Archetype: desk\_commercial; energy type: gas.

Figure 5.7 (cont.) Cambridge case study input demand profiles, grouped by sampling cluster, archetype, and energy type. Profiles have low opacity such that darker sections indicate significant profile overlap. See Table 5.3 for sampling cluster dates and Table 5.2 for buildings per archetype. ‘\_we’ = weekend, ‘\_wd’ = weekday.

## 5.3 Results

### 5.3.1 Scenario generation

Different methods were used to generate scenarios for the Bangalore and Cambridge case study districts. The Bangalore samples were generated prior to, and outside the scope of this thesis, using fPCA (Ward *et al.*, 2018). In this chapter, the samples were applied to case study buildings by scaling them to building floor area and randomly assigning them to days corresponding to the pre-defined monthly sampling clusters.

For Cambridge, the acquired data included several buildings describing each archetype. Consequently, KDE was applied to multi-building sets. 96 PDFs are required, to describe every combination of energy type (2), archetype (4), and sampling cluster (12). To make the most of the limited observations existing for buildings in the Cambridge district, 5-fold cross-validation was used to calculate the best-fit bandwidth in the range  $(0, 1]$  with either the ‘gaussian’ or ‘top hat’ (i.e. uniform distribution) kernel, using the Python package Scikit-learn (Pedregosa *et al.*, 2011) for KDE and Hyperopt (Bergstra *et al.*, 2013) for minimising the error between the validation data and the PDF. Table 5.4 shows the resulting bandwidths and kernels following 5-fold cross-validation. Only six PDFs are described using a top-hat kernel, with the rest using a gaussian kernel. The bandwidth of the kernels varies from 0.01 to 0.49. The four highest bandwidths are used alongside a top-hat kernel, which is to be expected, as a top-hat kernel ends more abruptly than a gaussian kernel. A higher bandwidth allows the observations to run more smoothly into each other, but the choice of top-hat kernel ensures that the probability density function ends abruptly close to the lowest and highest observations.

	Heat				Electricity			
	commercial desk	lab	research desk	lab	commercial desk	lab	research desk	lab
Term1_wd	0.04997	0.05887	0.20264	0.12394	0.08637	0.03682	0.03647	0.0373
Term1_we	0.03101	0.08895	0.08343	0.057	0.01124	0.04261	0.03714	0.03247
Vac1_wd	0.1158	0.08505	0.07766	0.07911	0.0483	0.29988	0.09826	0.03743
Vac1_we	0.13675	0.1211	0.10433	0.05748	0.05955	0.07582	0.06453	0.05468
Term2_wd	0.49099	0.11374	0.07045	0.26323	0.04754	0.381	0.05114	0.14478
Term2_we	0.0457	0.02587	0.04204	0.03974	0.11849	0.01128	0.04833	0.02955
Vac2_wd	0.04018	0.06603	0.08173	0.04574	0.39096	0.02128	0.02703	0.05492
Vac2_we	0.01451	0.04203	0.02651	0.02466	0.01844	0.04253	0.03385	0.035
Term3_wd	0.05666	0.19518	0.04705	0.12393	0.06037	0.07404	0.04446	0.058
Term3_we	0.02685	0.04919	0.05434	0.08039	0.02702	0.03195	0.07398	0.03055
Vac3_wd	0.05348	0.12213	0.07176	0.04954	0.0512	0.07614	0.15696	0.05651
Vac3_we	0.07554	0.0764	0.08497	0.07413	0.02375	0.04223	0.06033	0.03264

Table 5.4 Calculated bandwidths for each KDE subset in the Cambridge case study input dataset. Underlined values refer to those selected for use with a ‘top-hat’ kernel; all others use a ‘gaussian’ kernel. Background bars highlight relative magnitude of bandwidth.

Figure 5.8 gives the range of all the 500 scenarios of district-wide energy demand per utility for both case studies. Mean demand profiles are shown as lines and minimum to maximum range of any scenario is given as a shaded region (also marked on the right side of the y-axis). Each scenario is an aggregation of the demand for each building and every day in the year.

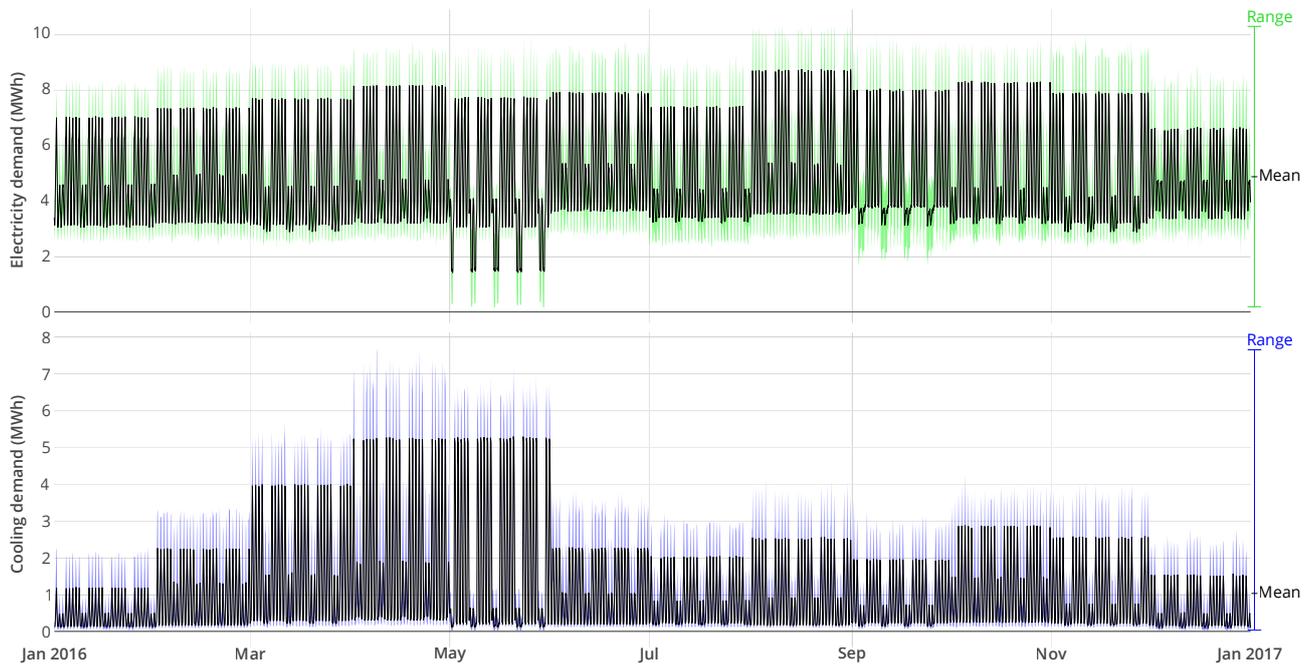
The demand in Bangalore is described with 24 distinct sampling clusters (weekend and weekday per month) and the demand in Cambridge is described with 4 archetype buildings, each with 12 sampling clusters (academic term times and vacation). These sampling clusters are evident in the samples, particularly between weekdays and weekends. Electricity demand shows less variation throughout the year than thermal demand. Indeed, the variation in thermal demand is not sufficiently well described by the choice of sampling clusters. A greater number of sampling clusters would better represent the data, leading to smoother transitions in both case studies. But a greater number of sampling clusters would lead to fewer observations available when producing PDFs.

There is a high possible deviation from the mean in any of the 500 scenarios, particularly for electricity and heat demand in Cambridge (Figure 5.9b). The demand in any time interval could range from 0.5x to 1.5x the mean. Within a set of days associated with the same sampling cluster, there is much less variation in the peak demand. This variation remains below 1MW in the mean curve and 5MW in the min/max range. As the profiles in Figures 5.8 and 5.9 aggregate all buildings in the district, the deviation of the profile mean and min/max is lower than that which is visible on a building-level.

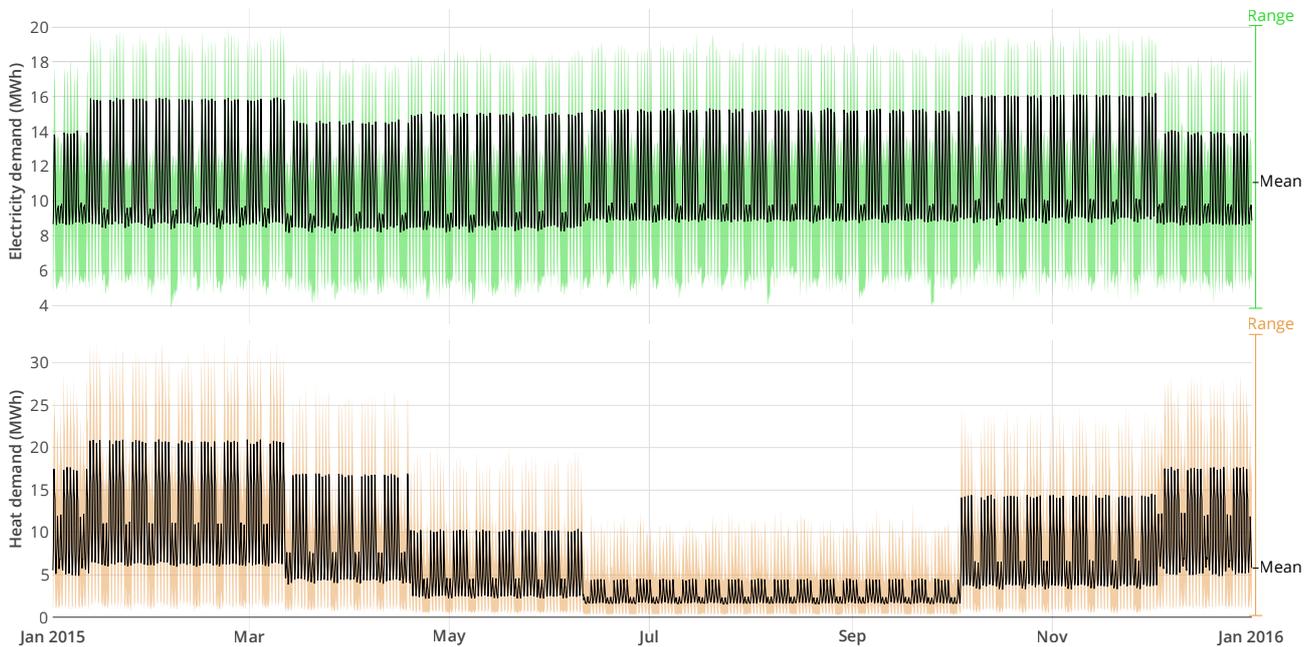
The mean annual demand of the modelled Cambridge district, as determined by the average of all 500 scenarios, is approximately 1.61x the amount predicted in the energy masterplan. The modelled mean annual demand is 142GWh electricity and 68.2GWh heat compared to 88GWh and 42GWh respectively predicted by the masterplan. The masterplan is based on a single archetypal annual demand mapped to all buildings of that archetype. As such, it is possibly less accurate than the estimate from SG, suggesting that a greater demand can be expected on the site. Neither the result of SG nor that of the masterplan can be validated, however, as none of the buildings on the masterplan site have a detailed physics-led model associated with them. If they did, only once commissioned in approximately ten years' time would results be truly validateable. In Bangalore, the demand cannot be compared to a district prediction, as no masterplan exists for this illustrative study.

As aforementioned, the optimisation is run only over typical days (TDs), not the full year given in Figure 5.8. Clustering based on the sampling clusters (24 for Bangalore, 12 for Cambridge) would have been the most straightforward approach to selecting TDs from the generated scenarios. However, the length of the resulting timeseries was too long for optimisation tractability in both case studies. Instead, extreme day masking followed by k-means clustering was undertaken independently on each of the 500 scenarios. For Cambridge, four extreme days were selected from the timeseries: maximum difference between PV available resource and electricity demand, maximum difference between ST available resource and thermal demand, maximum thermal demand, and maximum electricity demand. For Bangalore, the same masking was undertaken, but ignoring the ST mask, since that technology is not present in the model. The remaining, unmasked days were then clustered into 9 days for Bangalore and 2 days for Cambridge. In total, 12 and 6 TDs were used in the optimisation model for Bangalore and Cambridge, respectively.

The impact of just using TDs, compared to using the full timeseries, can be calculated ex ante. The variation between the full scenario demand timeseries and the clustered timeseries



(a) Bangalore, India.



(b) Cambridge, UK.

Figure 5.8 Total hourly district demand in one typical year, as sampled for 500 scenarios. Mean profiles are depicted by the darker area. To ensure clarity, the profile of all 500 scenarios is not displayed, instead the minimum to maximum range of all scenarios is given as a lighter shaded region.

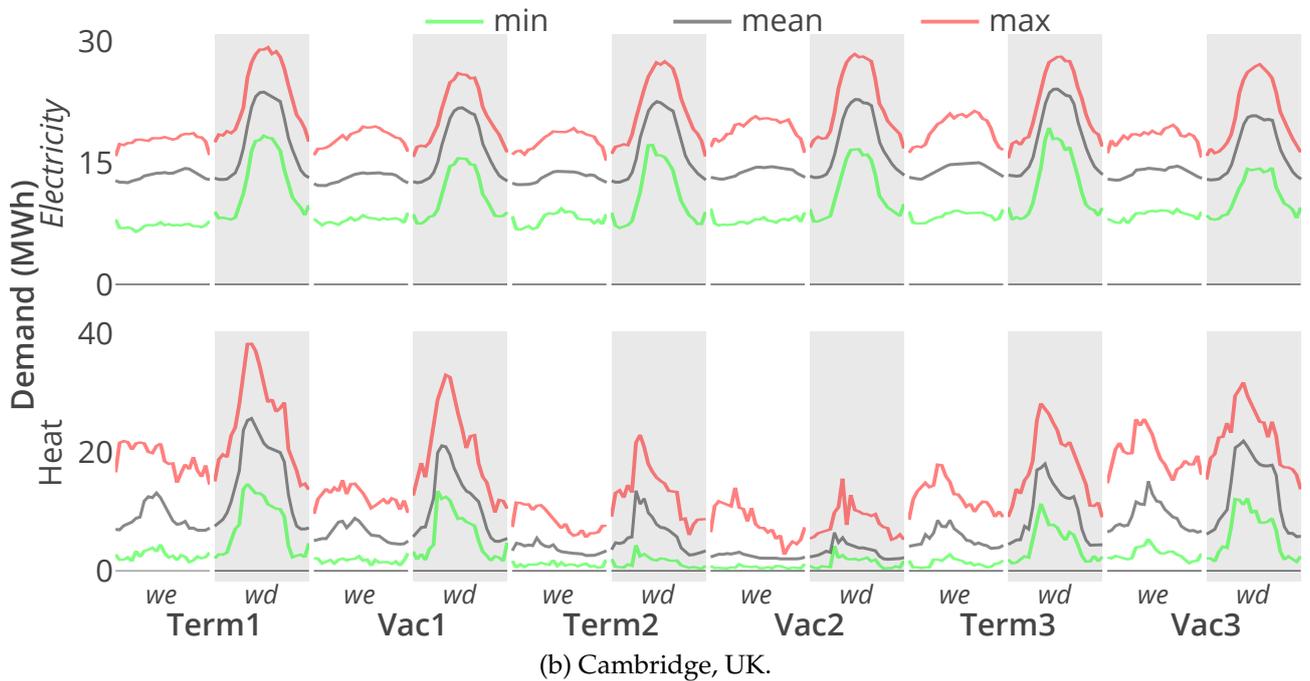
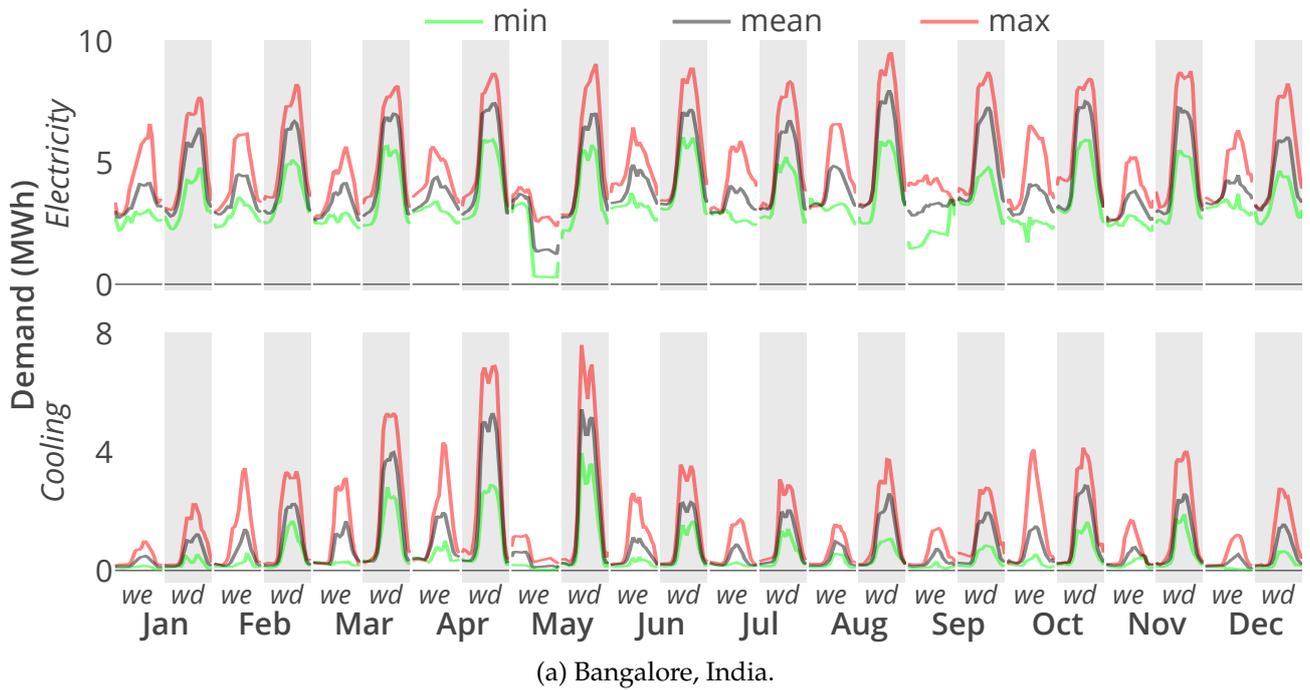


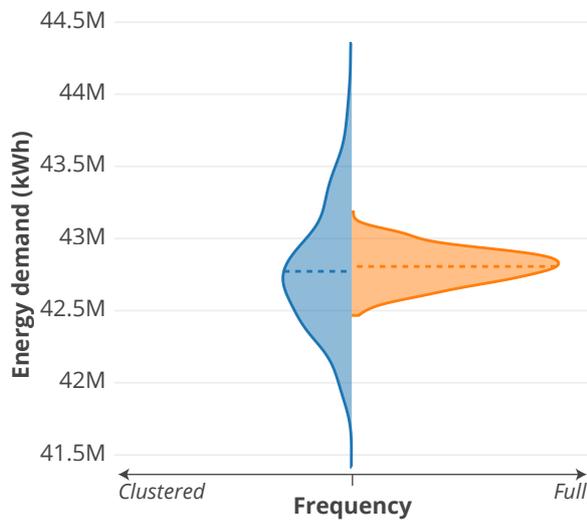
Figure 5.9 TD demand profiles for both Bangalore and Cambridge case study districts. Three traces are given per TD. The mean trace shows the mean demand across all 500 scenarios and all days within the TD group. Maximum and minimum traces are the extreme demand profiles from any of the 500 scenarios, on any of the days within the TD group.

is shown in Figure 5.10. The mean of electricity demand is well captured in the Bangalore case, but all other distributions show a clear mismatch created when clustering. A less conservative demand estimate is made in each case, more so when clustering thermal demand. In fact, only two clustered scenarios have annual cooling demand greater than the lowest full timeseries scenario cooling demand in the Bangalore case. For the purposes of ensuring demand is met, it is useful to be conservative with demand, but the result of automated clustering shows that the opposite is the case. Thus, it is clear that the objective function value will be underestimated. This expectation is in agreement with previous studies concerned with the impact of time clustering (Kotzur *et al.*, 2018a; Pfenninger, 2017a). Clustering is a necessity in this chapter, due to the introduction of the additional scenario dimension. To aid in improving optimisation results, intra-cluster storage management is implemented<sup>5</sup>.

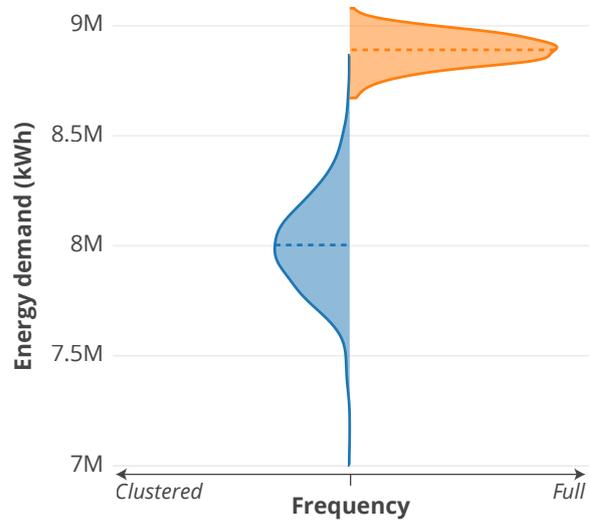
In addition to the ex-ante impact of using TDs, a Bangalore demand scenario was optimised with increasing timeseries aggregation, to assess the impact on the result. Scenario number 70 was chosen, since the sum of annual demand in that scenario is close to the average of annual demand across all scenarios. The full year demand profile for scenario 70 was clustered by first masking the same three extreme days as previously discussed: maximum difference between PV resource and electricity demand, maximum cooling demand, and maximum electricity demand. The remaining 363 days were then clustered by k-means into a different number of TDs, from 1 to 256. The optimisation results from each of these models can be compared, to understand the impact of reducing the time dimension. Figure 5.11 shows that there is a constant reduction in PV investment and increase in ECh investment as the number of TDs approaches the full timeseries. Yet, to realise this difference, a two order of magnitude increase in solution time is required. Table 5.5 shows that the solution time increases consistently with increasing number of TDs, except at 67 and 131 TDs, where the solution times appear to be reversed: the latter takes half as long to solve as the former. The additional solution time is caused by optimising the linear relaxation model. The 131 TDs model is also anomalous as it has the lowest optimal investment cost of any number of the models, although its total system cost (i.e. objective function value (OFV)) fits within the sequence as expected. The OFV increases with increased number of TDs, caused by increased operation costs. This is the expected result, given that the clustered demand results given in Figure 5.10b are shown to be less conservative.

---

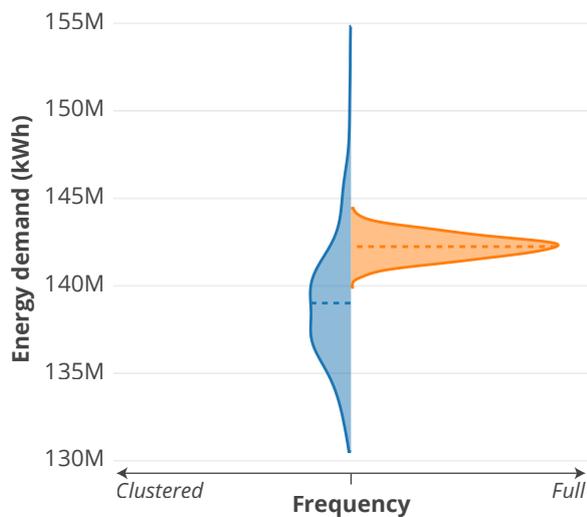
<sup>5</sup>Inter-cluster storage constraints have been implemented as per Kotzur *et al.* (2018b).



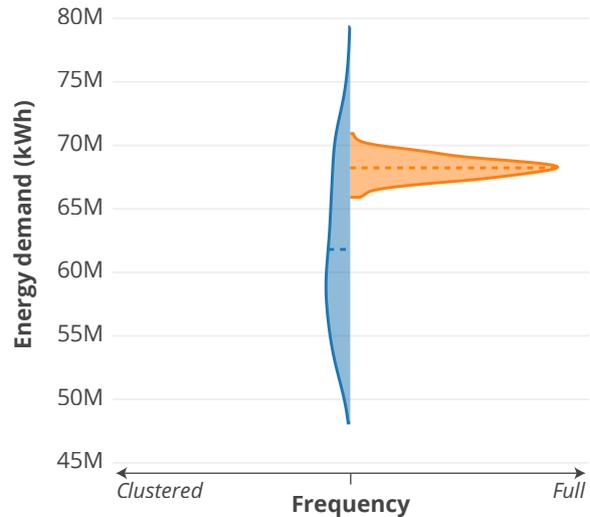
(a) Electricity; Bangalore, India.



(b) Cooling; Bangalore, India.



(c) Electricity; Cambridge, UK.



(d) Heat; Cambridge, UK.

Figure 5.10 Comparison of annual energy demand distributions between that derived from the full timeseries and from the clustered timeseries of 500 generated scenarios in Cambridge and Bangalore. The clustered timeseries consists of six TDs and 12 TDs in the Cambridge and Bangalore case studies, respectively. Frequency of annual energy demand derived from the clustered timeseries is shown as positive in the negative  $x$  direction, while the frequency of annual energy demand derived from the full timeseries is positive in the positive  $x$  direction.

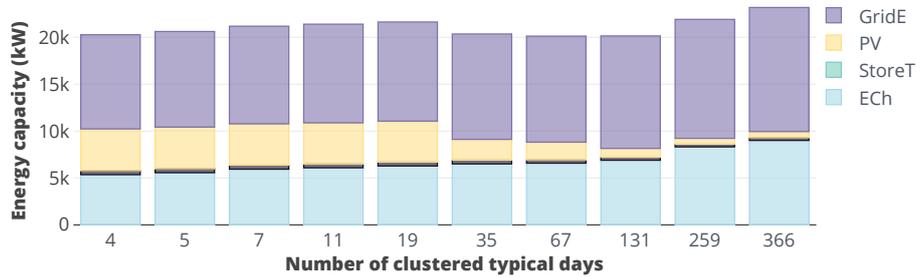


Figure 5.11 Technology energy capacity in the cost minimised result of the Bangalore case study model, using scenario 70 demand profiles and increasingly aggressive (right to left) timeseries aggregation. TDs were selected by first masking three extreme days in every model, followed by k-means clustering of the remaining days, to give a total of 4 to 259 TDs. Energy capacities resulting from optimisation of the full, unclustered timeseries (366 days) are also shown.

Table 5.5 Comparison of solution time, total system costs, and annual energy demand associated with the optimal solution of the Bangalore case study scenario 70 model, following increasingly aggressive (right to left) timeseries aggregation. Three days were masked in every model, with the remaining days being clustered into 1 to 256 TDs.

	Number of model TDs									
	4	5	7	11	19	35	67	131	259	366
<b>Solution time</b>	11	13	17	29	70	288	1017	547	1472	3134
<b>Cost</b>										
Operation	29.1	29.3	29.8	30.2	30.2	33.2	33.3	34.7	34.9	35.7
Investment	4.69	4.74	4.84	4.87	4.92	3.59	3.40	2.88	3.12	3.36
<i>Sum</i>	33.8	34.0	34.7	35.1	35.1	36.8	36.7	37.6	38.0	39.0
<b>Demand</b>										
Electricity	41.9	42.1	42.7	43.1	42.9	42.8	42.4	42.4	42.0	42.8
Cooling	7.27	7.53	8.02	8.16	8.36	8.60	8.59	8.30	8.27	8.95
<i>Sum</i>	49.2	49.6	50.7	51.3	51.3	51.4	50.9	50.7	50.3	51.8

### 5.3.2 Scenario reduction

When discussing the methodology in section 5.1.2, two approaches to scenario reduction were introduced. Figure 5.12 allows us to compare two metrics describing the 500 scenarios: annual total system demand and independently optimal OFV. The better the correlation between the two metrics, the more likely the simpler of the two (annual total system demand) could be used for scenario reduction. Clustered annual demand varies in Cambridge by 24%, compared to an OFV variation of 16% (Figure 5.12a). There is a trend for higher system cost with an increase in total system demand. In fact, the highest and lowest OFV is associated with the highest and lowest system demand, respectively. However, the OFV distribution is skewed towards lower values, leading to a long tail towards high values. This skew is not exhibited in the total system demand. Additionally, the fourth highest OFV has a total system demand lower than more than 30 other scenarios. In Bangalore, there is a skew in both demand and OFV (Figure 5.12b). However, the highest demand scenario results in the 15<sup>th</sup> highest OFV. There is a smaller percentage variation in both total system demand and OFV compared to the Cambridge data, at 7% and 11% respectively.

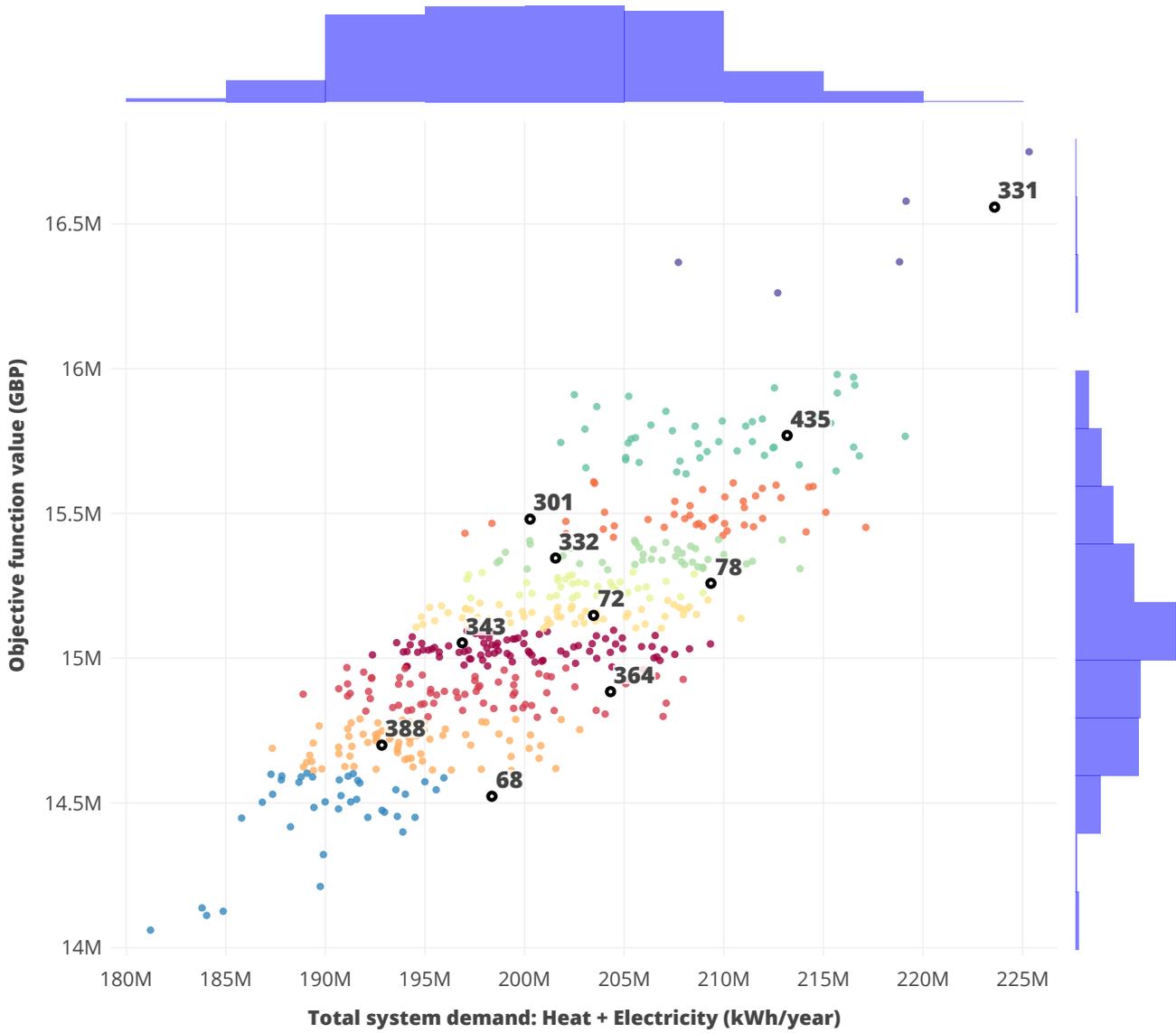
In either case it might be reasonable to select reduced scenarios based on annual demands, instead of independently optimal OFVs, but differences would be evident at the extremes of the data. Given the heavy influence these extreme scenarios are likely to have on the scenario optimisation OFV, selecting based on independently optimal OFV is prudent in both case studies. Solution times vary across each independent run, 100 – 400 seconds for Cambridge models and 25 – 50 seconds for Bangalore models. Given that these models can be run in parallel, there is little reason to avoid this step.

The ten reduced scenarios chosen to represent the full set are also highlighted in Figure 5.12. For Cambridge, there is a clear subset of six scenarios which have a higher OFV than the rest of the scenario set. This leads to the highest reduced scenario (331) being the 2<sup>nd</sup> highest total system demand scenario and 3<sup>rd</sup> highest OFV scenario. In Bangalore, the corresponding scenario (6) is not so close to the extreme scenarios. In fact, the highest OFV scenario is 3% higher than the scenario chosen to represent it.

Table 5.6 shows that the maximum thermal and electricity demand is not given by the same scenario in Bangalore, although it is in Cambridge. The proportion of scenarios a single reduced scenario represents ranges from 1.2% to 18.8%; this corresponds to the weight of each reduced scenario in the objective function. The lower weight reduced scenarios are, as expected, those at the extremes of OFV. Thus, costs incurred in an extreme scenario will have approximately 10 times less influence on the OFV than in a moderate scenario (see Equation 5.1). In the optimisation, this translates to a risk-based compromise, where one is willing to accept a relatively high operational cost in unlikely futures (the extreme scenarios) as they have an order of magnitude lower impact on OFV.

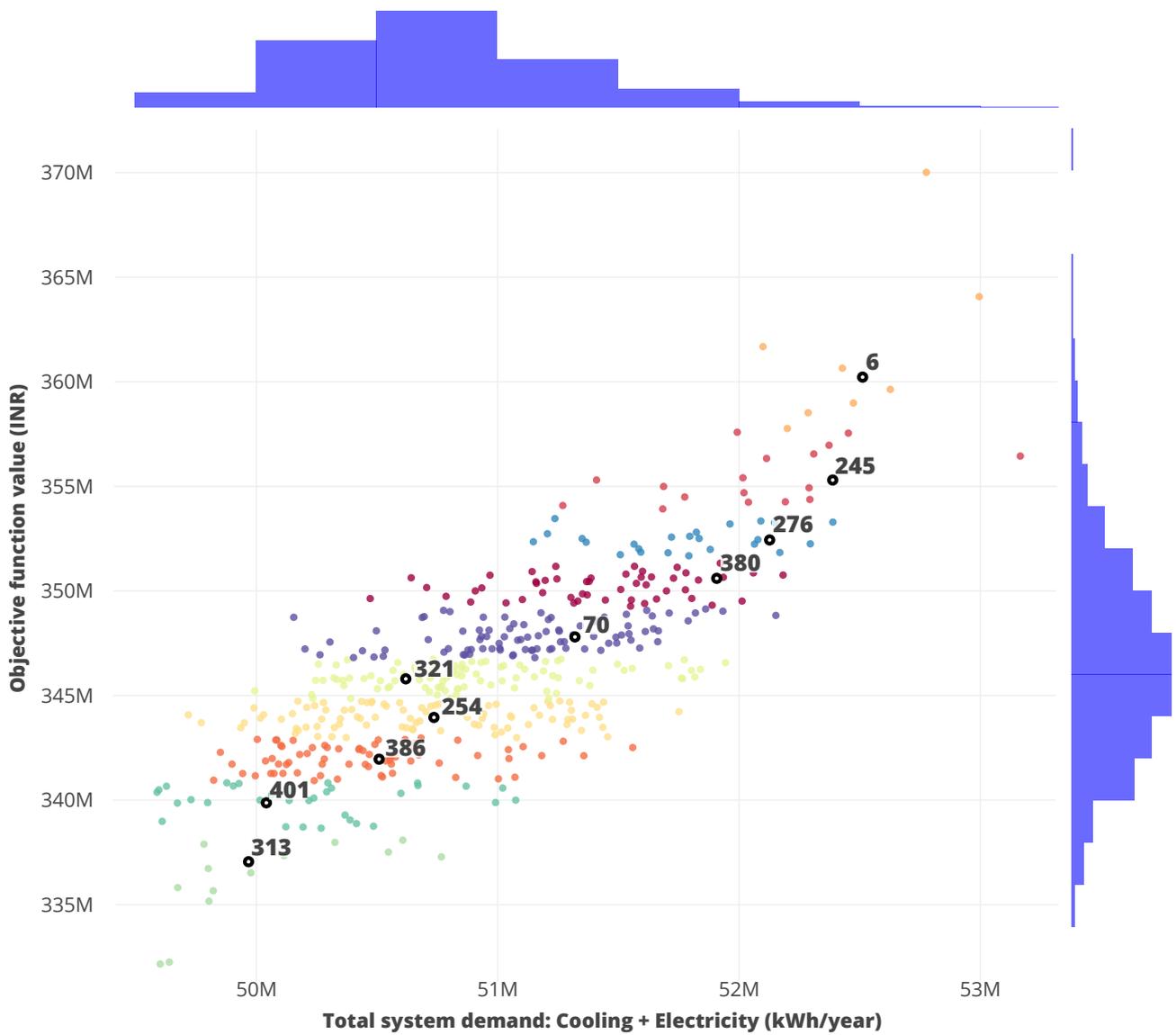
### 5.3.3 Scenario optimisation

In this section, the result of SO is compared with the result from running a single scenario model. The single scenario is taken to be the mean of the 500 generated scenarios, to emulate



(a) Cambridge, UK.

Figure 5.12 Total district demand compared to independently optimal OFV for 500 demand scenarios. Ten scenarios chosen to represent the full set following scenario reduction are highlighted and numbered. Scatter points are coloured based on the scenario by which they are represented. The distribution of demand and OFV is given outside the scatter plot.



(b) Bangalore, India.

Figure 5.12 (cont.) Total district demand compared to independently optimal **OFV** for 500 demand scenarios. Ten scenarios chosen to represent the full set following scenario reduction are highlighted and numbered. Scatter points are coloured based on the scenario by which they are represented. The distribution of demand and **OFV** is given outside the scatter plot.

Table 5.6 Description of 10 reduced scenarios chosen to represent the full set of 500 generated scenarios, given to three significant figures. Scenarios are ordered by **OFV** (left to right ascending), as this is the value on which **SR** takes place. Highest values in each category are highlighted in .

Scenario	68	388	364	343	72	78	332	301	435	331
<b>Demand</b> (x10 <sup>7</sup> kWh)										
Heat	6.68	5.64	7.01	5.68	6.47	7.15	5.93	5.52	6.92	7.23
Electricity	13.2	13.6	13.4	14.0	13.9	13.8	14.2	14.5	14.4	15.1
Sum	19.8	19.3	20.4	19.7	20.3	20.9	20.2	20.0	21.3	22.4
<b>OFV (x10<sup>7</sup>GBP)</b>	1.45	1.47	1.49	1.51	1.52	1.53	1.54	1.55	1.58	1.66
<b>Weight (%)</b>	8.00	15.2	14.0	16.6	11.4	7.40	8.60	8.60	9.00	1.20

(a) Cambridge, UK.

Scenario	313	401	386	254	321	70	380	276	245	6
<b>Demand</b> (x10 <sup>6</sup> kWh)										
Cooling	7.64	7.96	8.07	7.89	7.91	8.17	8.31	8.31	8.23	8.30
Electricity	42.3	42.1	42.4	42.7	42.6	42.9	43.3	43.6	44.0	44.0
Sum	50.0	50.0	50.4	50.6	50.5	51.0	51.6	51.9	52.2	52.4
<b>OFV (x10<sup>8</sup>INR)</b>	3.37	3.40	3.42	3.44	3.46	3.49	3.52	3.54	3.56	3.61
<b>Weight (%)</b>	3.00	7.00	14.0	18.2	18.8	18.2	10.4	5.00	3.60	1.80

(b) Bangalore, India.

the use of ‘archetypal’ profiles of demand. Running the Cambridge ‘mean’ model took 238 seconds to solve with an **OFV** of  $1.52 \times 10^7$  GBP. The Bangalore ‘mean’ model took 45 seconds to solve with an **OFV** of  $3.39 \times 10^8$  INR. The Cambridge case took longer to solve due to its larger spatial size; it is four times larger than the Bangalore model.

### Risk-neutral

Optimisation of the 10 reduced scenarios by risk-neutral **SO** took 30 minutes and 4.32 hours to reach a solution for the Bangalore and Cambridge cases, respectively. This constitutes a two order of magnitude increase in solution time relative to the respective mean models. The results follow similar trends in both case studies, as shown by the installed capacity at each node given in Figure 5.13. Only demand nodes have installed capacity; there are no energy centre technologies included in the investment portfolio. This is caused by the prohibitive cost of laying thermal network pipes, a cost not often included in district network studies. Renewable energy technologies are also not considered in the Cambridge case.

Storage technologies are introduced in the Cambridge **SO** result, alongside an increase in both national grid and boiler capacity (Figure 5.13a). No new technologies are incorporated in the Bangalore case, but there is a similar trend towards increasing conventional energy technology capacity (Figure 5.13b). However, **PV** capacity reduces by almost 1MW in the Bangalore **SO** result; this is the only capacity reduction caused by incorporating uncertainty.

Figure 5.14 shows that due to the smaller installed capacity in the mean model, conventional technologies will run more frequently. The capacity factor (total carrier production as a proportion of installed energy capacity) of conventional technologies is particularly high in the mean model, even greater than any of the 500 independently optimal scenarios for GridE, ECh, and NGB. This leads to less available capacity for meeting additional demand in the mean model. Conversely, SO technologies have greater capacity and lower capacity factor. The range of capacity factor for SO conventional technologies differs to that of the independent scenarios, being lower in the Bangalore case and greater in the Cambridge case.

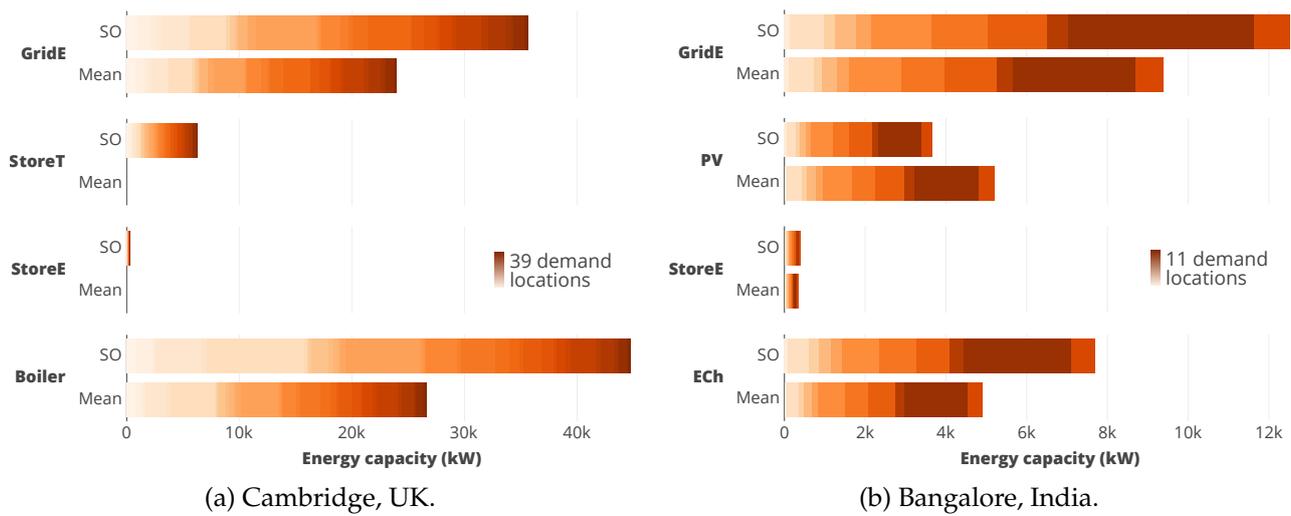


Figure 5.13 Installed capacity of technologies to achieve the optimal OFV in both mean (single scenario) and SO cases. Available technologies with zero capacity have been omitted, such as all those at the energy centre. The contribution from each node to the total technology capacity has been differentiated with a colour gradient. Although, in some cases (e.g. 'StoreT' in (a)), not all nodes have installed capacity.

Incurred costs are high as a result of SO. In Cambridge, there is an additional investment cost of  $+2 \times 10^5$  GBP required to meet the demand of 10 SO scenarios. However, operation cost realised in any of the 10 scenarios is low. The mean operation cost reduces relative to the mean model and to the mean of the 500 independent scenarios (Figure 5.15a). This is particularly important, as the operation costs contribute to over 95% of the overall costs. In Bangalore, the costs are inverted: investment cost is marginally lower than the mean model, but operation costs are significantly higher (Figure 5.15b). This is caused by the use of load shedding in low probability scenarios, rather than increase technology capacity. This is somewhat counterintuitive, since operation costs are approximately six times greater than the investment costs. In both cases, the mean OFV (sum of operation and investment costs) is higher in SO than in the mean model or the independently optimal models. This does not make the single scenario models better, as the cost they portray relies on the demand being exactly that given by the single scenario. They cannot necessarily meet demand in all future scenarios used in SO, let alone doing so at a lower cost.

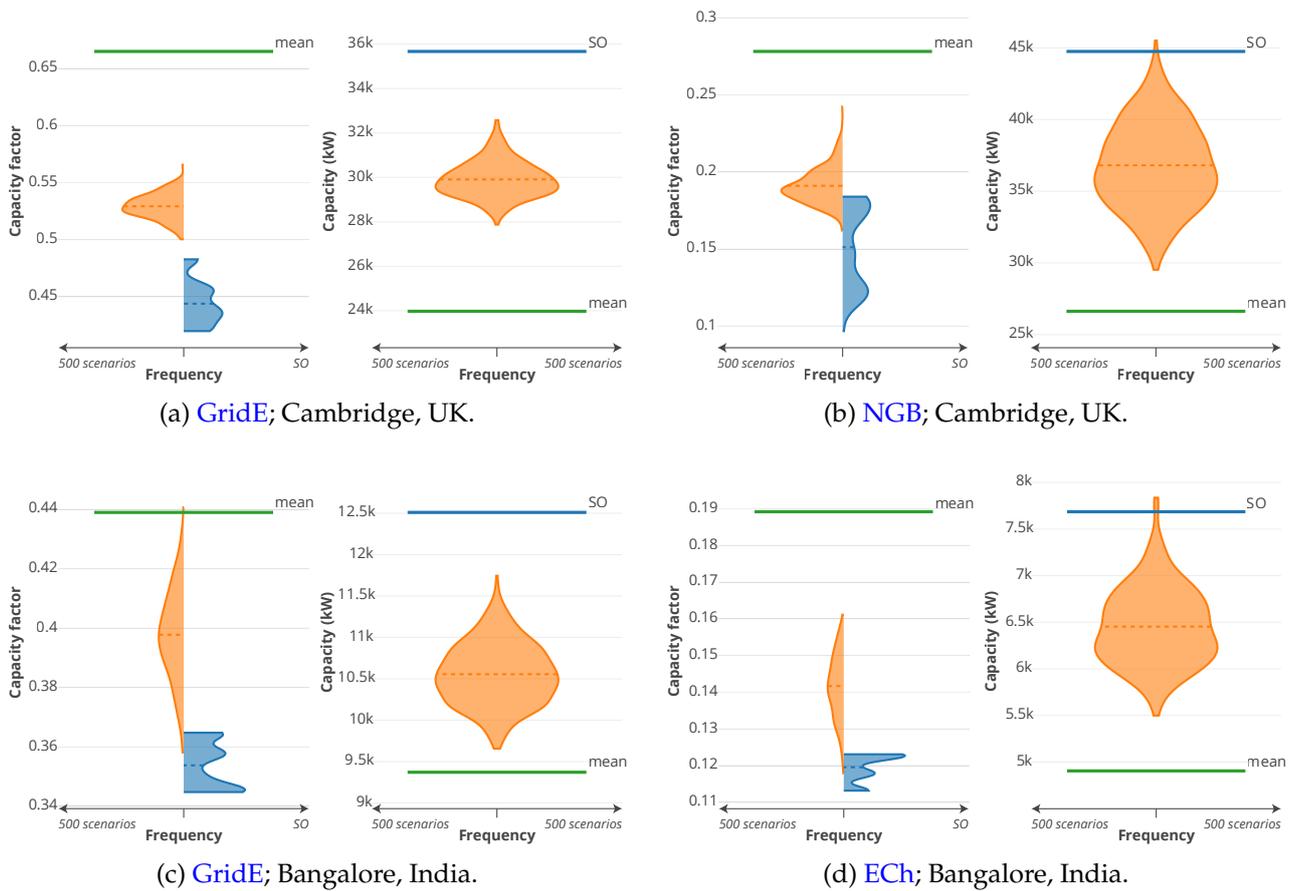
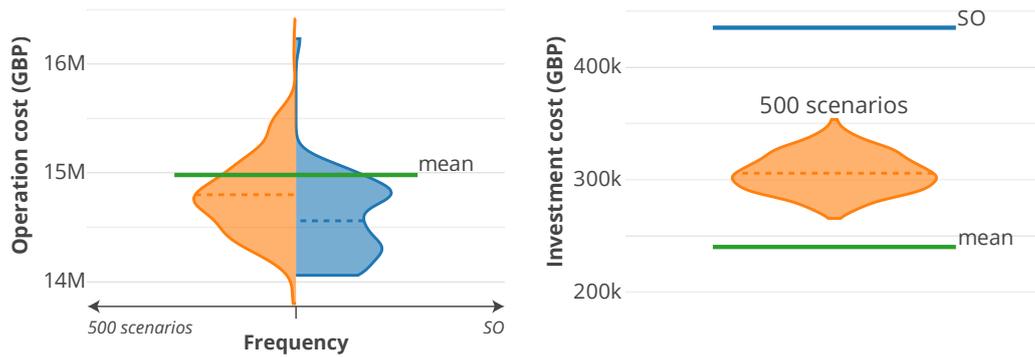
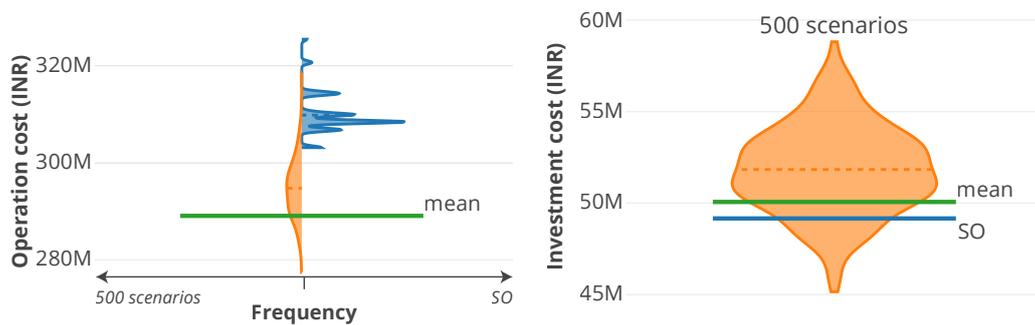


Figure 5.14 Installed capacity (right) and capacity factor (left) for conventional technologies chosen to achieve the optimal OFV in the mean, 10-scenario SO, and 500 independent scenario model runs for both Cambridge and Bangalore case studies. Distributions are given for SO capacity factor (positive frequency in positive  $x$  direction), 500 independent scenario capacity factor (positive frequency in negative  $x$  direction) and 500 independent scenario energy capacity (frequency mirrored about the  $y$  axis). All other results are single values. Capacity factor is the total energy production / energy capacity.



(a) Cambridge, UK.



(b) Bangalore, India.

Figure 5.15 Contribution of investment and operation costs to the OFV of the mean model, 10-scenario SO model, and all 500 independent scenario models. Distributions are given for SO operation cost (positive frequency in positive  $x$  direction), 500 independent scenario operation cost (positive frequency in negative  $x$  direction) and 500 independent scenario investment cost (frequency mirrored about the  $y$  axis). All other results are single values.

## Risk-averse

Risk-aversion involves adding the CVaR term to the SO objective function (see Equation 5.2). In this section  $\alpha$  is set to 0.9, which will impact the 10% highest operation cost scenarios.  $\beta$ , the degree of risk aversion is set to 4 in the Bangalore case and to 16 in the Cambridge case; Section 5.3.3 explores the choice of  $\beta$  in more detail.

Figure 5.16 shows that the standard deviation of total system cost incurred on the realisation of a scenario is lower when including the risk-aversion term. This means a designer can be more confident of the expected system cost, but the mean cost is higher. Again, reducing risk leads to greater initial expectation of incurred costs. The change in shape of the system cost distribution is led by changes in operation cost: in Cambridge and Bangalore, the lowest operation cost increases. In Bangalore, the highest possible operation cost also decreases, thus mitigating the worst case cost for the same investment cost. In Cambridge, the worst case operation cost remains the same, but the investment cost reduces by 10%. This leads to the small reduction in maximum realisable cost seen in Figure 5.16a.

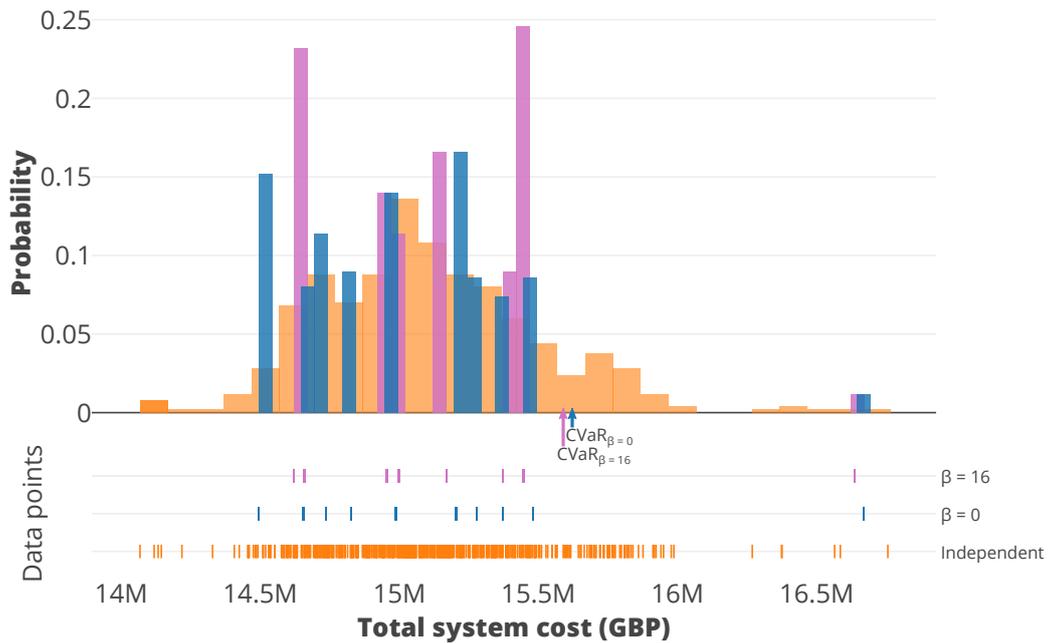
The CVaR has been lowered in both models, but only by a small amount. This matches the small variation in installed energy capacity seen in Figure 5.17. There is no perceptible difference between the risk-neutral and risk-averse capacities in the Bangalore case study, although the ECh capacity has marginally increased at the cost of reduced StoreE discharge capacity. The reduction in storage capacity is more evident in the Cambridge case study, where StoreT is more than halved and StoreE is removed completely. Although storage technologies offer flexibility in meeting fluctuations in demand, they are considered to be less cost effective than ensuring sufficient conventional technology capacity.

## Sensitivity analysis

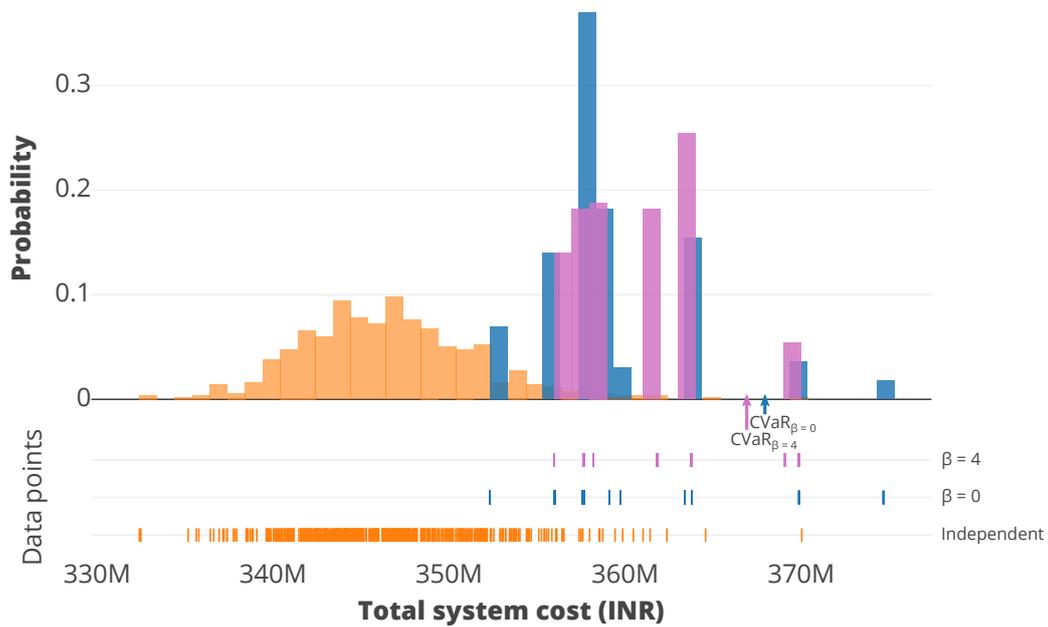
This chapter introduces a number of new hyperparameters to the optimisation problems, namely the number of reduced scenarios and the magnitude of risk aversion  $\beta$ . Unlike the breakpoint hyperparameters studied in Chapter 4, pre-optimisation of SO hyperparameters was not applicable. However, sensitivity analysis can be undertaken to assess the impact of hyperparameter variability.

The number of reduced scenarios and magnitude of risk aversion were varied in the Bangalore case study. Reduced scenarios were varied from 5 to 30, run each time against  $\beta$  values varying from 0 to 16. Figure 5.18 shows the change in the optimal system cost realised in each iteration of sensitivity analysis alongside solution time of each model. Three expected trends are exhibited. First, the range in realisable system costs increases as the number of scenarios considered in SO increases. This occurs as the weight associated with extreme scenarios decreases, allowing them to take on larger costs without greatly impacting the OFV. Second, as  $\beta$  increases, the maximum realisable cost scenario reduces. A larger  $\beta$  leads to greater risk aversion, which directly impacts the most extreme scenario(s). Third, the solution time increases as the number of scenarios (and thus number of decision variables) increases.

However, Figure 5.18 shows many instances in which exceptions to these trends occur. The range of realisable costs increases drastically at 20 scenarios (hence the need to separate Figures 5.18a and 5.18b), then decreases again at 30 scenarios. This extreme increase is caused



(a) Cambridge, UK.



(b) Bangalore, India.

Figure 5.16 Distribution of total system cost resulting from running models using risk-neutral SO, risk-averse SO, and risk-unaware optimisation on all 500 independent scenarios. Data points from which the distributions are constructed are identified. The CVaR resulting from running the risk-neutral and risk-averse SO models is also marked.

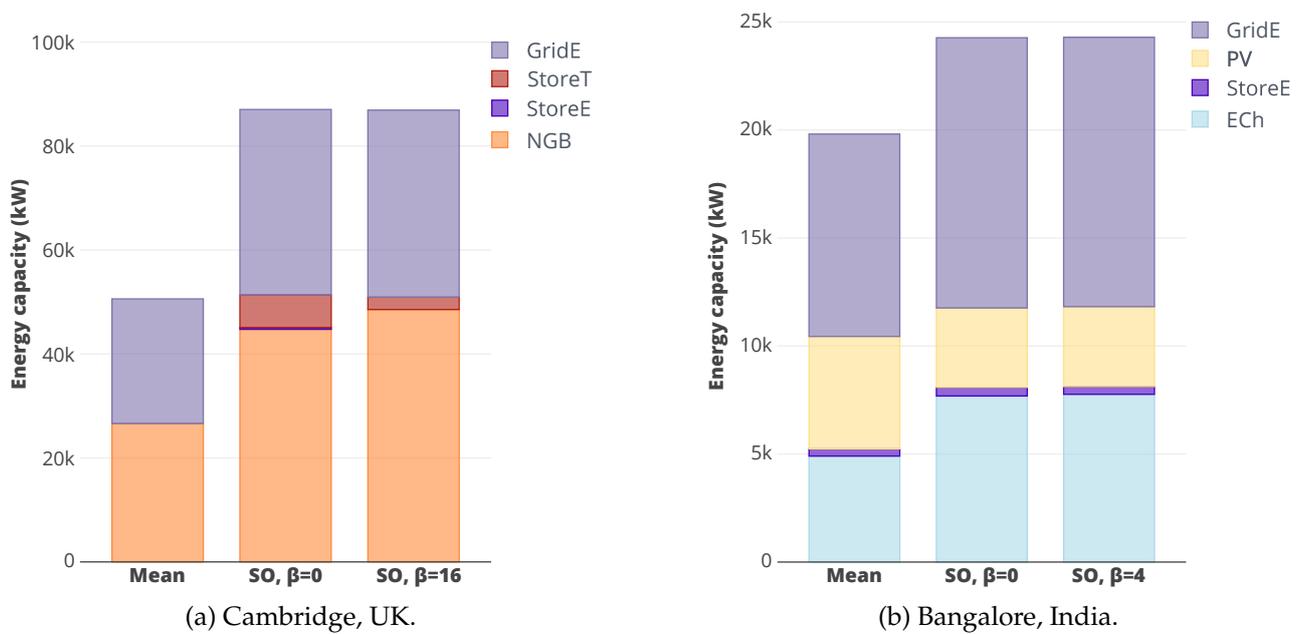
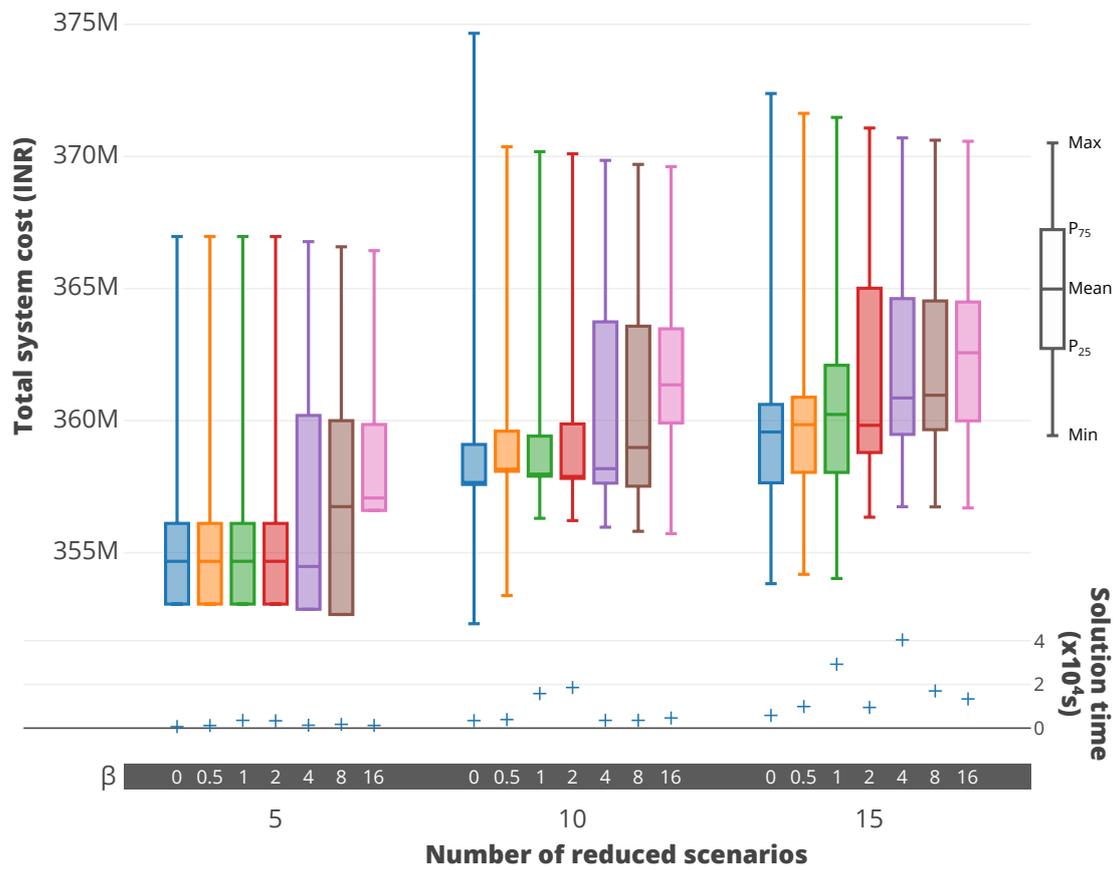


Figure 5.17 Optimal energy capacity resulting from running the mean model compared to using risk-neutral SO and risk-averse SO. Building-level technologies have been aggregated to give the total district-wide capacity for those technologies.

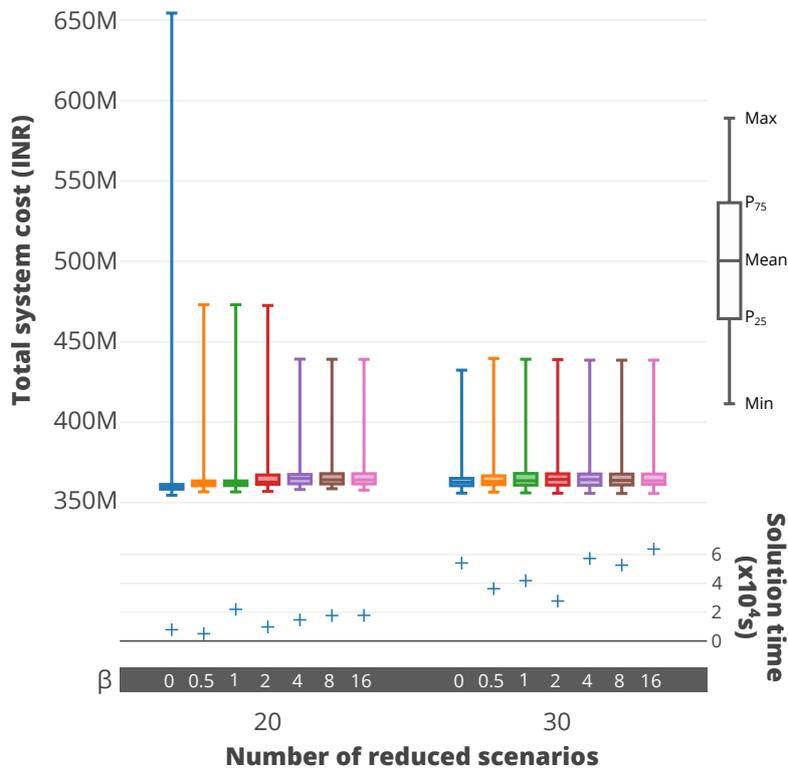
by the most extreme reduced scenario only representing itself. This gives it a weight of 0.2%, which is significantly lower than other scenarios in the reduced set. Excessive load shedding takes place in this one scenario, without noticeably impacting the OFV. The same scenario exists in the 30 reduced scenario set, but it no longer has a significantly lower weight than other scenarios in the set.

At 20 scenarios, the impact of  $\beta$  is most evident: the highest realisable cost reduces by a third. An additional 35 million INR is removed from this cost for  $\beta$  values above 2. At both 10 and 15 scenarios there is also a reduction in the most extreme cost following the introduction of  $\beta$ , which continues to decrease up to  $\beta = 16$ , albeit at a lower rate. At 5 scenarios, risk aversion only impacts the most extreme cost from  $\beta = 4$ . As there is only one scenario affected by the CVaR term at 5 scenarios (i.e. in the top decile of the cost PDF), this low impact is expected. The same cannot be said for the 30 scenario model, which has seven scenarios which are affected by the CVaR term. Yet, there is only a 1 million INR decrease from  $\beta = 0.5$  to  $\beta = 16$ . It is also the only model in which the highest realisable cost actually *increases* with the introduction of risk-aversion.

Solution time is a function of the number of decision variables, but some problems are more straightforward to solve than others. For instance no model run at 20 reduced scenarios takes longer than  $2 \times 10^4$  seconds to solve. At 15 scenarios, the solution time when  $\beta = 4$  is twice this limit, at  $4 \times 10^4$  seconds. The introduction of risk aversion leads to an iterative constraint, which requires compromising between the value of CVaR and the risk-neutral component of the objective function. If these components are particularly well-balanced, converging on the optimal solution will take time. Convergence issues are most evident between  $\beta = 1$ ,  $\beta = 2$  and  $\beta = 4$ . The solution time varies considerably between these levels



(a) 5 to 15 reduced scenarios.



(b) 20 and 30 reduced scenarios.

Figure 5.18 Result of sensitivity analysis on the levels of risk aversion  $\beta$  and number of reduced scenarios used in risk-averse SO, applied to the Bangalore case study. Comparisons are made between the range of system costs (investment + operational) and solution time following the optimisation of each model.

of risk aversion, with the direction of the solution time swing depending on the number of reduced scenarios.

Although there is a relatively large range in [OFV](#), this is not accompanied by a noticeable variation in investment decision. All [SO](#) models choose a similar combination of technologies to that given in [Figure 5.17b](#), which itself shows little variation between the risk-neutral and risk-averse investment capacity. Only at five scenarios is there a difference in investment capacity; both [GridE](#) and [ECh](#) have a 10% lower capacity than given in models using a greater number of reduced scenarios.

## 5.4 Out-of-sample testing

Although we expect to account for uncertainty by using **SO**, it is not immediately apparent how well a solution covers the possible realisations of the future. It is also not clear from just considering the investment portfolios resulting from **SO** how much *more* resilient a solution is to a particular risk compared to those objectives which did not consider that risk, or those that did not consider risk at all. Applying **Out-of-sample (OOS)** tests to the result of each objective in this study is one way of assessing resilience ex-post. Indeed, they are considered an effective approach to evaluate both decisions and the decision making tools used to generate them (Conejo *et al.*, 2010).

In the **OOS** tests applied in this section, the investment decisions which have already taken place, following either single scenario optimisation or **SO**. These investment decisions lead to fixed capacities which are then exposed to new operating conditions to evaluate their performance. The **OOS** scenarios providing the new operating conditions are stochastically generated demand scenarios, based on the scenarios generated in preparation for **SO**.

**OOS** performance is a measure of *robustness* or *graceful extensibility*, which cover two of the four concepts of resilience as defined by Woods (2015)<sup>6</sup>. A robust system is one which ‘expands the set of disturbances a system can respond to effectively’ (Woods, 2015). Invariance in a particular sense  $Z$  when perturbations of type  $Y$  are applied ensures robustness (Alderson and Doyle, 2010). Based on this definition, we would expect system robustness to have increased when accounting for uncertainty in system design, perhaps more so when risk aversion is incorporated. Where uncertainty is not considered in the initial investment optimisation, the same tests may be considered as a measure of graceful extensibility.

Unlike the definition of robustness, which measures invariance across a particular known uncertainty set, graceful extensibility is a measure of system brittleness (Woods, 2015). Indeed, even if variance in **OOS** tests can be kept low to ensure robustness, resulting systems can still fail catastrophically when exposed to the realisation of unknown uncertainty in system parameters (Read, 2005; Woods, 2015). The main goal of a gracefully extensible resilient system is to not be susceptible to catastrophic failure, even if this means great deviation from a pre-defined objective function. As such, if the **OOS** tests of single scenario results show a relatively low maximum magnitude of unmet demand, then they could be considered as meeting the requirements of graceful extensibility. If there is an unreasonably high quantity of unmet demand in the single scenario **OOS** tests which is significantly lowered on the introduction of **SO**, the latter design may instead be considered as having provided graceful extensibility. In the results pertaining to **OOS** tests, the variance and magnitude of unmet demand can be separated to better understand the measure of resilience [2] and [3]. Low variance indicates a robust system, while a low maximum magnitude indicates graceful extensibility.

---

<sup>6</sup>Woods (2015) also defines two other concepts of resilience, robustness as rebound (resilience [1]) and sustained adaptability (resilience [4]). Neither is deemed as relevant to the optimisation approach considered in this research. Resilience [1] cannot be understood, since there is little inter-timestep connection, i.e. only storage technologies retain knowledge of other timesteps. Resilience [2] requires a greater extent to the boundary of the system, particularly the interconnection of the energy system with other networks over longer time periods.

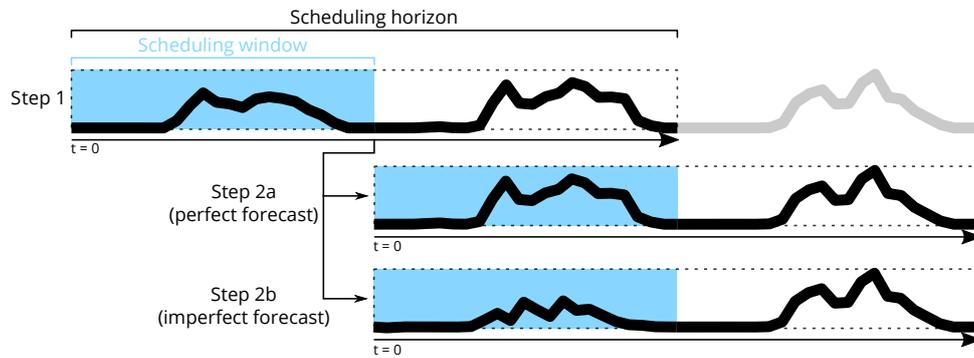


Figure 5.19 Depiction of a rolling horizon optimisation, where decisions are made in step 1 for the scheduling window, considering an optimisation over the scheduling horizon. Step 2 constitutes a step forward in time, and another optimisation. Steps 2a and 2b show the realisation of a perfect and imperfect forecast, respectively.

### Rolling horizon optimisation

When running an energy system model to make investment decisions, there is perfect knowledge of the parameters in every time step, even if there are multiple scenarios. By fixing technology capacities based on the optimal technology portfolio, a model can be run using a rolling horizon instead. Although optimisation will take place for the full time series, it is done in smaller chunks of a few days; each chunk is the scheduling horizon. Optimisation is run over each horizon, with the intent of implementing only a sub-period: the scheduling window. The period between the end of the window and the end of the horizon is considered a ‘forecast’. It is only used to provide a trajectory for the scheduling window. As can be seen in Figure 5.19, the horizon will ultimately become the window, as the optimisation steps through time. With perfect foresight (Figure 5.19, step 2a), the parameter values do not change when moving from the scheduling horizon to the scheduling window. Instead, to emulate imperfect forecasting, the horizon can be given different parameter values to those seen in the window, for the same time steps (Figure 5.19, step 2b). Existing studies have considered both perfect and imperfect forecasting. Where uncertainty has been introduced, it has included pricing (Jin *et al.*, 2018; Khodabakhsh and Sirouspour, 2016), weather (Jin *et al.*, 2018; Verrilli *et al.*, 2017), and demand (Giraud *et al.*, 2017; Khodabakhsh and Sirouspour, 2016; Kopanos *et al.*, 2013). In the perfect foresight case, there is redundancy in the optimisation; only the results from the scheduling window are actually retained. However, the small timeseries of a few days can still reduce overall computation time, as linear programming algorithms tend towards solving in a polynomial time, relative to their size (Megiddo, 1986).

In this chapter, the scheduling horizon and window are 24 hours and 12 hours, respectively. Perfect foresight is assumed in the demand uncertainty in OOS tests, for which there are 500 scenarios, as produced in the SG step prior to SO (Section 5.1.1).

### 5.4.1 Robustness to unexpected demand

SO leads to a more robust technology portfolio, in both Cambridge (Figure 5.20a) and Bangalore (Figure 5.20b). By considering 10 scenarios in SO, a much greater range of futures have been accounted for in the system design, constituting a two order of magnitude reduction in unmet demand in any realisable scenario compared to the mean model. The maximum OFV and annual demand independent scenarios also compare favourably to the mean model, reducing unmet demand by an order of magnitude, although neither manages to improve on the SO result.

Unmet demand exists in every OOS test, no matter what the initial investment portfolio. Even when running the corresponding OOS scenario as that which led to the investment decisions, there is unmet demand. For instance, the highest OFV/demand scenario in the Cambridge case (346) has 0.76MWh and 1MWh of unmet electricity and heat demand, respectively, when run in rolling horizon mode against the same scenario demand data. Not only does rolling horizon mode use data from the full timeseries, thus exposing the inadequacies in timeseries aggregation, but it also does not allow for storage planning beyond a two day time horizon. The impact of timeseries aggregation on the highest OFV scenario in the Bangalore case study (419) is shown in Figure 5.22. The impact is stark: a two order of magnitude reduction in unmet demand is realised when testing a single-scenario full timeseries model, compared to the same tests applied when using 12 TDs. In fact, with unmet demand not increasing above  $6 \times 10^2$  for either energy type, the single scenario full timeseries model performs considerably better than even the risk-neutral SO model. It can thus be expected that running the SO model using the full timeseries information would yield even greater improvements. However, for the models in this chapter it is not tractable to do so.

Greater robustness is realised by use of risk aversion. Figure 5.21 shows that this reduction is relatively small and only perceptibly impacts unmet electricity demand in Bangalore and unmet heat demand in Cambridge. The reduction in unmet electricity demand in Bangalore is particularly surprising, given the similarities between risk-neutral and risk-averse energy capacity (Figure 5.17b). The reduction in unmet heat demand in Cambridge can be explained more clearly, as risk-aversion led to the reduction in StoreT capacity. Operating a storage device is more difficult when no information is available beyond the 48 hour time horizon.

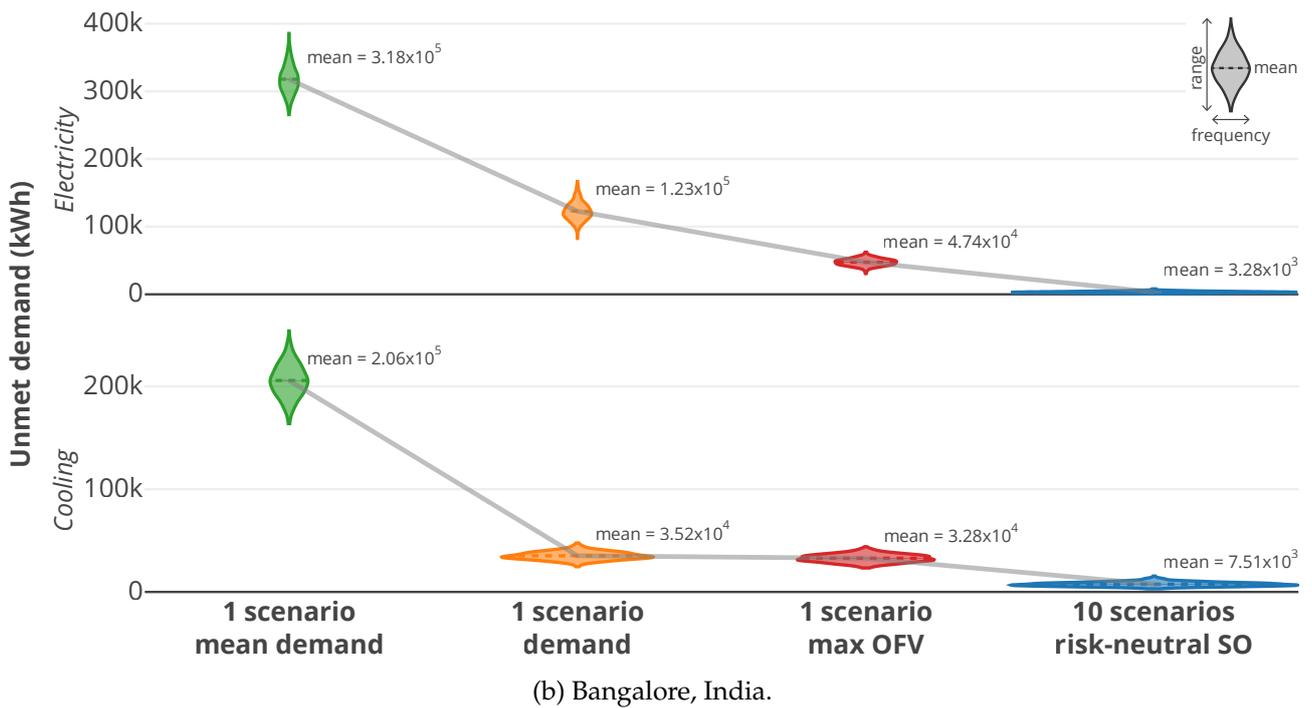
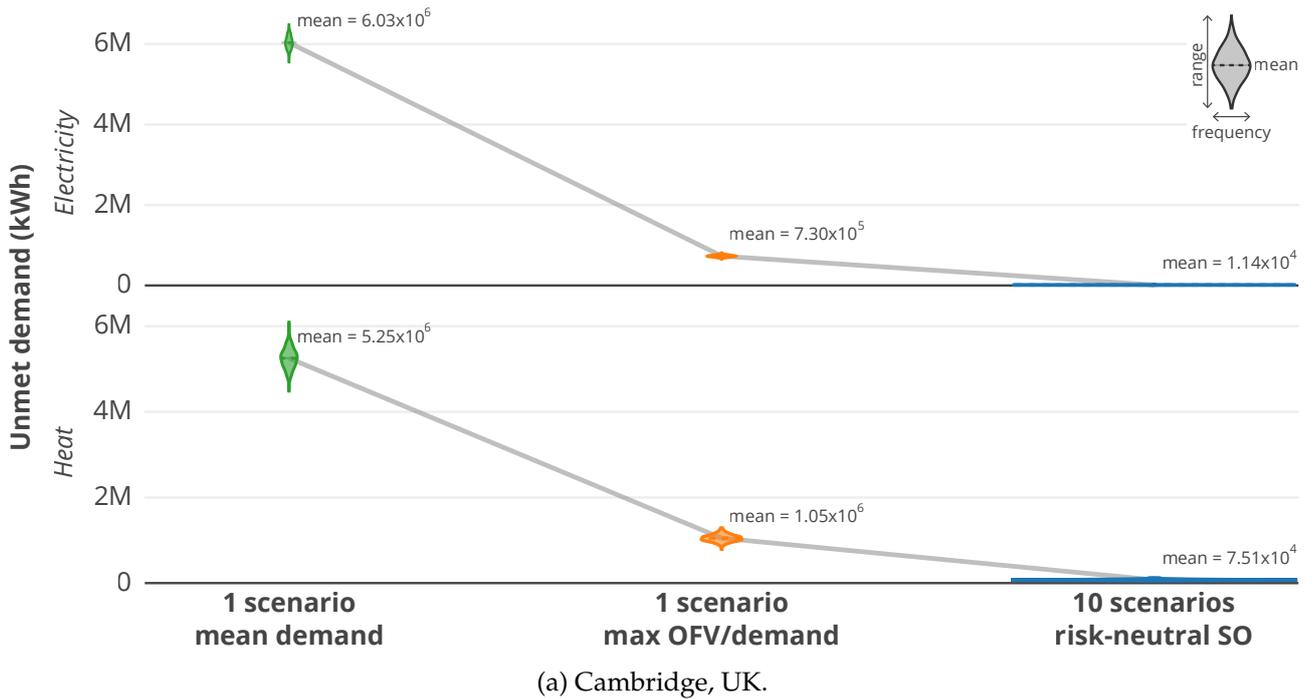


Figure 5.20 System unmet thermal and electricity demand distributions when running OOS optimisation tests in both Cambridge and Bangalore case studies. The 500 OOS scenarios are those resulting from SG. Technology capacities have been fixed in the tests to the result of optimising the mean scenario, scenario with highest OFV (when optimised independently), scenario with the highest total annual demand, and 10-scenario SO. The same scenario has highest OFV and total annual demand in the Cambridge case study. Distributions are given as violin plots, whereby size on the y axis give the range of demand values and size on the x axis gives the frequency of occurrence.

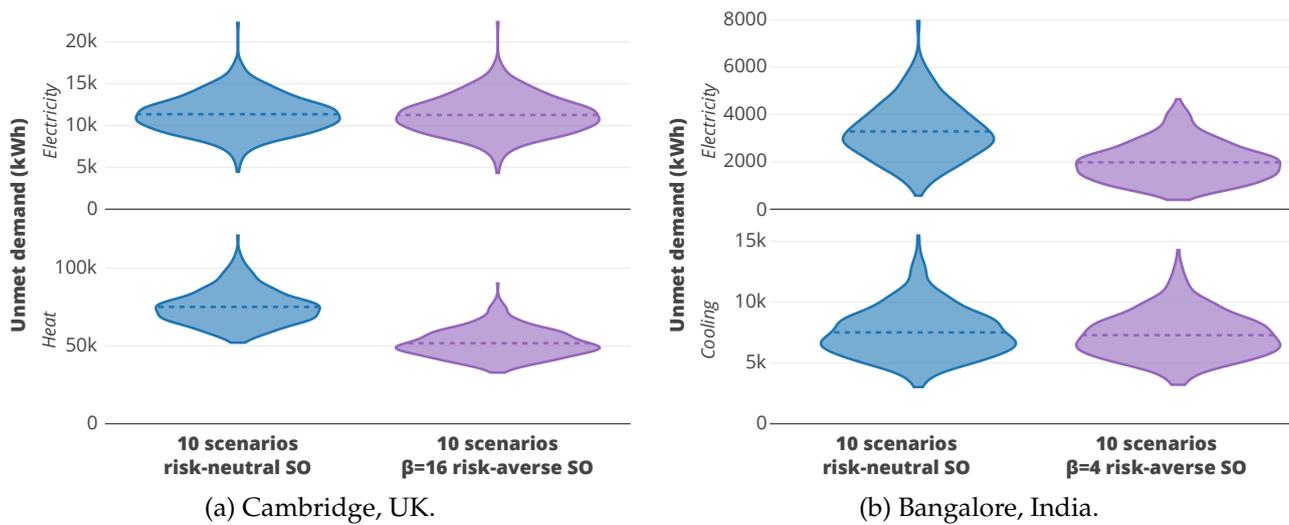


Figure 5.21 System unmet thermal and electricity demand distributions when running OOS optimisation tests in both Cambridge and Bangalore case studies. The 500 OOS scenarios are those resulting from SG. Technology capacities have been fixed in the tests to the result of optimising the 10 reduced scenarios in SO with and without risk-aversion. Distributions are given as violin plots, whereby size on the  $y$  axis give the range of demand values and size on the  $x$  axis gives the frequency of occurrence.

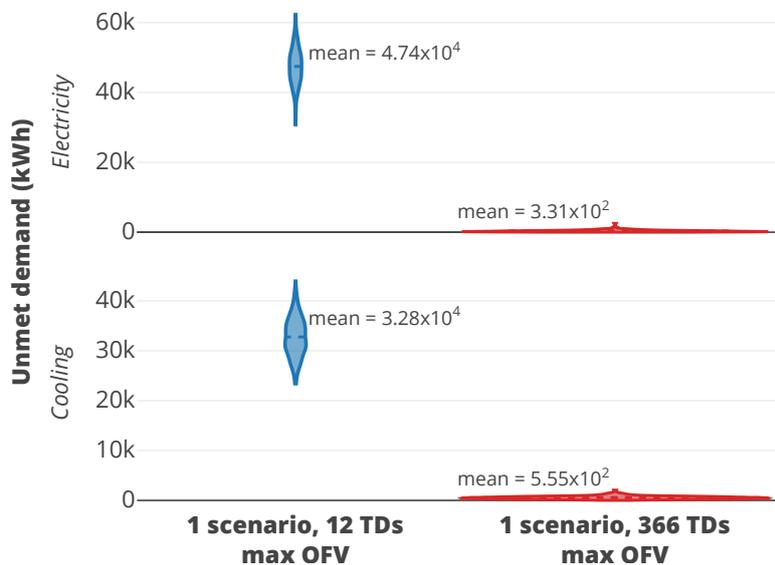


Figure 5.22 System unmet cooling and electricity demand distributions when running OOS optimisation tests for the same demand scenario in the Bangalore case study with either no timeseries aggregation or aggregation to 12 TDs. The chosen scenario was that which gave the highest OFV when optimised independently. Distributions are given as violin plots, whereby size on the  $y$  axis give the range of demand values and size on the  $x$  axis gives the frequency of occurrence.

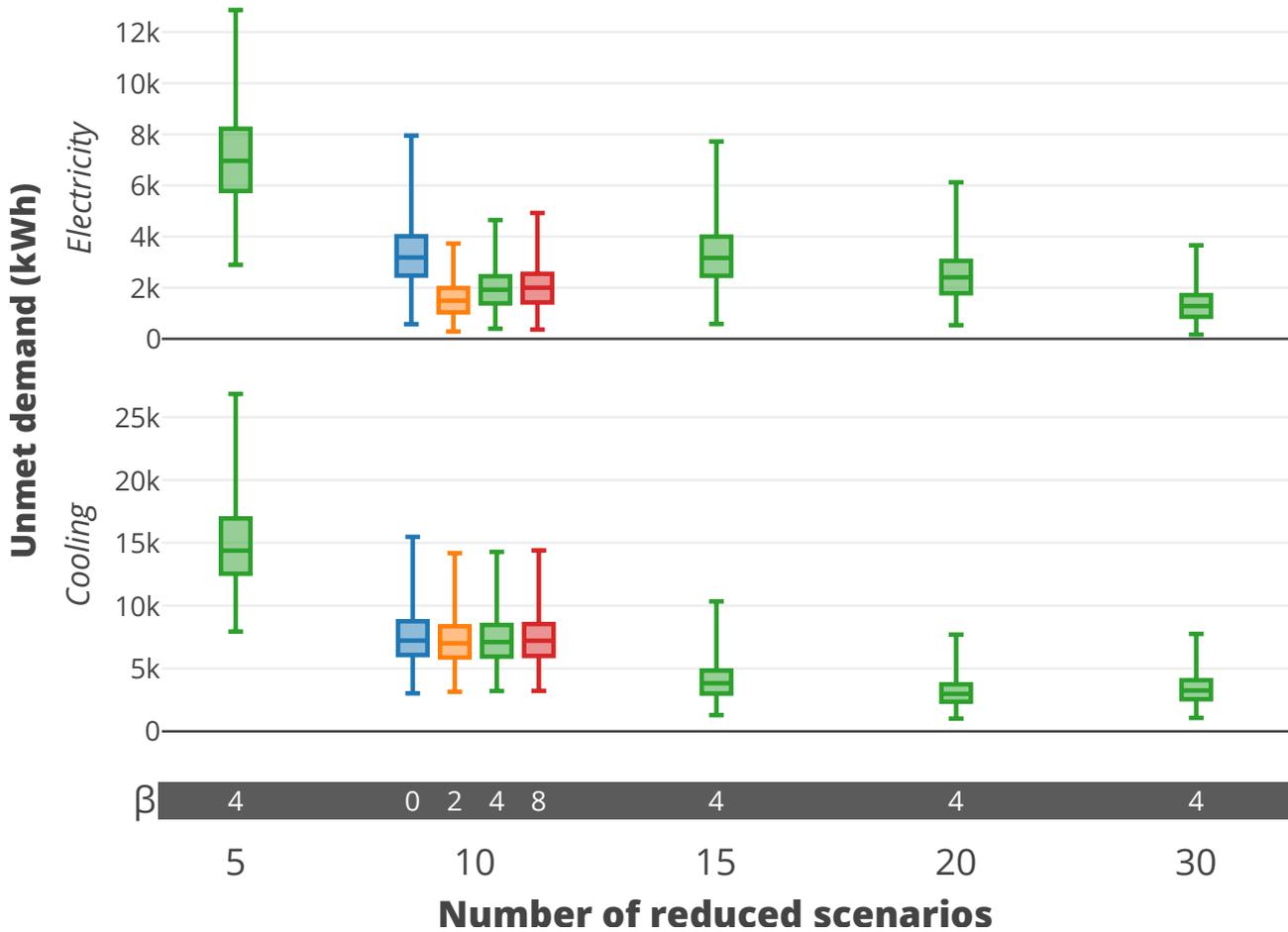


Figure 5.23 Results of sensitivity analysis on the unmet cooling and electricity demand realised in OOS optimisation tests for varying levels of risk aversion  $\beta$  and number of reduced scenarios, applied to the Bangalore case study.

### Sensitivity analysis

OOS tests have been run on a subset of the results of the sensitivity analysis given in Section 5.3.3, including  $\beta = 2$  to  $\beta = 8$  for 10 scenarios and  $\beta = 4$  for 5 to 30 scenarios (Figure 5.23). Increasing the number of scenarios in the reduced scenario subset reduces unmet demand, although less unmet electricity demand is realised when using 10 reduced scenarios instead of 15 or 20. Equally counterintuitive is the increase in unmet electricity demand with increasing risk aversion at 10 reduced scenarios. Although the introduction of risk aversion has clearly had a positive impact, heavily penalising high cost scenarios is less useful.

## 5.5 Impact of power interruptions

### 5.5.1 Reliability of national grid electricity

Power supply interruptions provide shocks to an energy system which can lead to catastrophic failure of energy supply. The majority of district energy system studies, which concentrate on European case studies, do not need to model with a high probability of these system shocks. However, many areas of Bangalore experience almost daily, unexpected interruptions to national grid power supply, caused by load-shedding. In fact, the local utility provider for Bangalore maintains an unscheduled power outage database<sup>7</sup>. Analysing this database, it is apparent that there is a great deal of uncertainty in how many distinct interruptions will take place, when they will occur, and for how long.

Using the recorded power outage periods from December 2014 to July 2015 in Bangalore, we have constructed a two-part probabilistic representation of intermittency. First, for every day in the year, there is a 70% chance of an outage. Second, for those days randomly selected as having an outage, the start and duration of the morning period (00:00 to 11:59) and the afternoon period (12:00 to 23:59) outages are randomly sampled from lognormal distributions describing the input data. An outage in a twelve hour period  $n$  can encroach on the next twelve hour period,  $n + 1$ , if the start time and duration lead to that. However, a sampled outage may not continue into the period  $n + 2$ , as that suggests sampling from a distribution tail which does not exist in the data.

The distribution of a full year of sampled intermittency is given in Figure 5.24. As can be seen, there is a greater than 50% chance that an interruptions will occur around 10:00 and 19:00 – 20:00. The probability reduces drastically overnight, and is somewhat lower in the middle of the day. This profile of interruptions matches the shape of electricity consumption in the state, as periods of high consumption are more likely to lead to load-shedding, if the infrastructure in place is inadequate to handle such levels of demand. Indeed, the sampled commercial office electricity demand used in this study, described in Figure 5.6, shows a similar profile shape, particularly on weekdays.

### 5.5.2 Impact of power interruptions on unsuspecting systems

Unreliable access to GridE is used to augment the OOS tests applied in Section 5.4. Particularly, given the unexpected nature of power interruptions, an imperfect horizon has been applied to the rolling horizon method. As such, only in the scheduling window are power outages realised, in the latter 12 hours of the scheduling horizon, access to the grid is assumed to again be reliable.

Across the 500 OOS tests, the same 366-day scenario of stochastically generated interruptions is used. Although several scenarios for intermittency in the year could be created, it would not be as easy to compare OOS test results. However, demand will still be varied according to the 500 generated scenarios given in Section 5.3.1.

<sup>7</sup><https://www.bescom.org/upo/public.php>

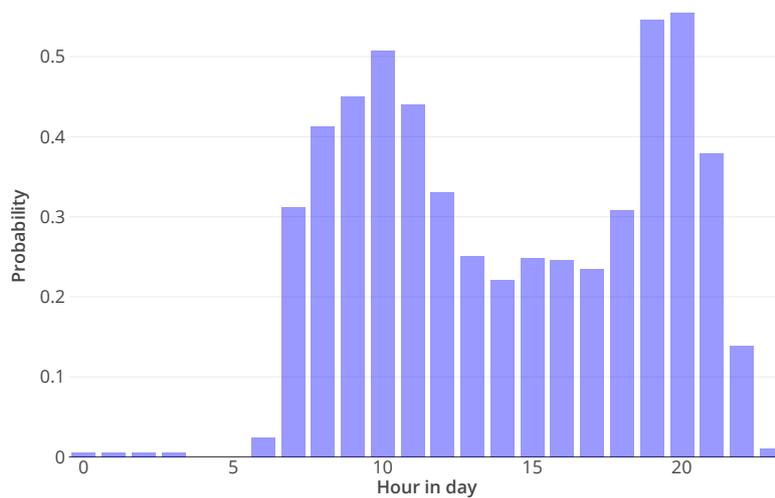


Figure 5.24 Distribution of sampled power interruptions across the Bangalore case study year, based on historical unscheduled interruption data in Bangalore. Each bar represents the probability of there being no electricity available from the grid in a given hour of the day.

When the models introduced in Section 5.3 are subjected to power interruptions, it is clear that they do not gracefully extend to manage this shock. Figure 5.25 shows that the unmet electricity demand is similar between the mean and SO models, both risk-neutral and risk-averse. Unmet demand has increased by two and four orders of magnitude for the mean and SO models, respectively, compared to the reliable grid OOS tests. Cooling demand is unaffected, however. Although cooling is met electrically using an ECh, this demand is met in preference to building electricity demand.

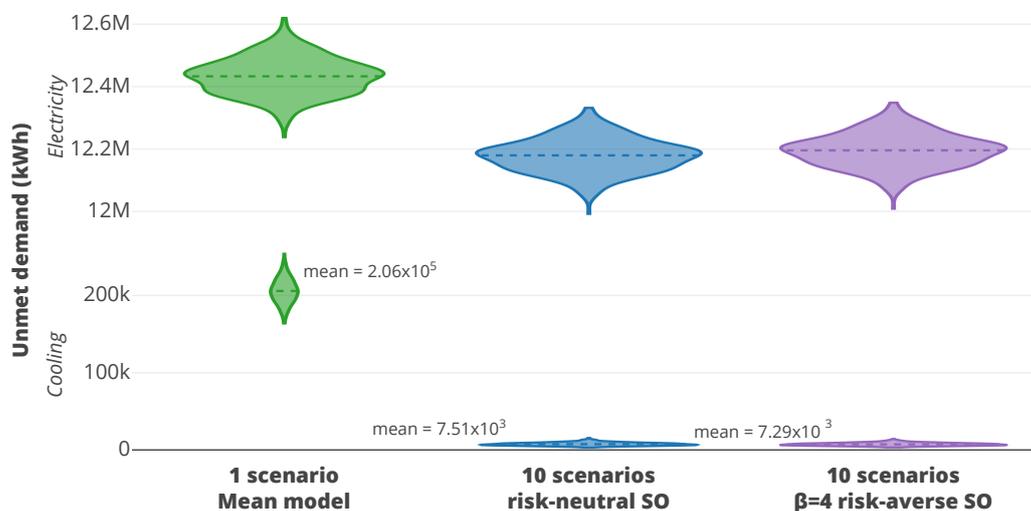


Figure 5.25 System unmet cooling and electricity demand distributions when running OOS optimisation tests which include electricity power supply interruptions, applied to the Bangalore case study models. Distributions are given as violin plots, whereby size on the  $y$  axis give the range of demand values and size on the  $x$  axis gives the frequency of occurrence.

### 5.5.3 Introducing resilience

Resilience to power interruptions can be incorporated into the model in several ways. In the first instance, the system can be designed with the expectation for some degree of expected

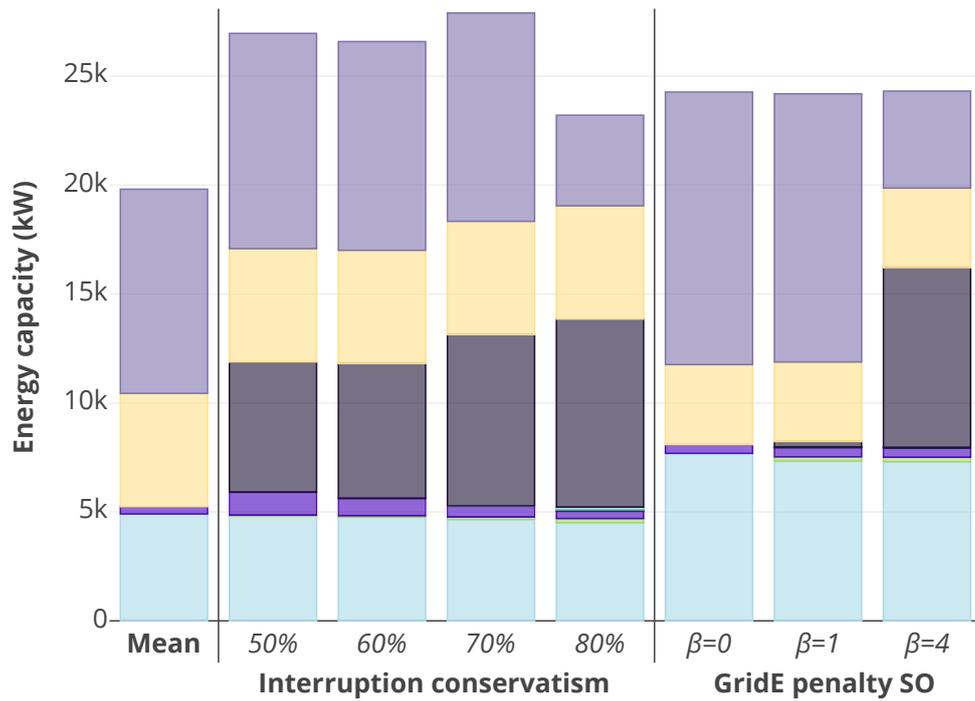
power interruption. Electricity supply availability is considered as a timeseries parameter in the mean model, with select hours in each day set to zero availability. Using the hourly power interruption probability seen in Figure 5.24, four levels of availability can be considered. The most conservative level involves setting any hour with an interruption probability greater than 20% to zero availability (80% conservatism). The least conservative level considers only those hours with an interruption probability greater than 50% to zero availability (50% conservatism). The two levels in between have 60% and 70% conservatism. Rather than explicitly applying zero electricity availability, it is also possible to update the CVaR constraint in the SO model to penalise only the dependence of the system on grid supply. In doing so, a financial incentive is applied to not using the electricity grid. As it is given in the CVaR term, it will most affect the high dependence scenarios.

Applying electricity supply conservatism to the mean model leads to investment in building-level DGs to meet demand (Figure 5.26a). Yet, the electricity produced by the DGs only exceeds that purchased from the grid when conservatism reaches 80% (Figure 5.26b). Investment in a biomass fuelled CCHP (B-CCHP) also occurs, but its contribution is consistently low. PV capacity does not change relative to the mean model as it is already the highest possible capacity for the available roof area. At low conservatism, there is greater use of StoreE, but this decreases by two thirds up to 80% conservatism.

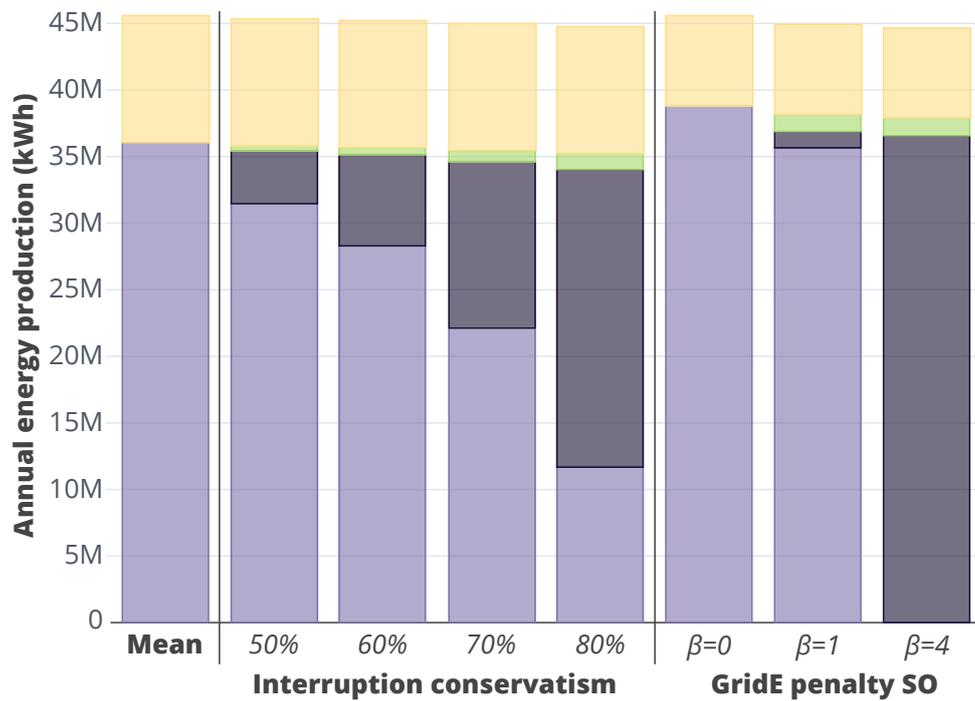
By updating the SO model to penalise only high dependence on the grid, the system quickly becomes autarkic, i.e. it no longer depends on any electricity from the grid. Although there is supposedly some GridE capacity at a risk aversion level of  $\beta = 4$  (Figure 5.26a), actual purchases from the grid are negligible (Figure 5.26b). Instead, building-level DGs are the optimal choice. The invested capacity in DGs is similar to 70% grid availability conservatism, but the subsequent production from that investment is three times greater. If  $\beta = 1$ , grid dependency remains high, with less than 10% of the risk-neutral SO grid dependence transferred to DG or B-CCHP production.

The impact on OOS tests is perhaps expected: there is a reduction in unmet demand when incorporating power interruption resilience. The impact is exhibited on both electricity and cooling demand, with the possibility of reducing both by several orders of magnitude. Figure 5.27 compares only the models which have incorporated power interruption resilience, in which there is a two order of magnitude reduction in both unmet electricity and cooling demand between 50% conservatism and GridE risk-averse SO. As the magnitude of error decreases, the magnitude of the range of error also decreases, meaning the expected unmet demand is more certain. Additionally, using the electrically autarkic  $\beta = 4$  GridE risk-averse SO results, a lower unmet cooling demand is realised as compared to the original  $\beta = 4$  risk-averse SO model result shown in Figure 5.21.

An additional 23 million INR of investment is required to install the most conservative systems, whether resulting from single scenario conservatism or GridE risk-averse SO. This corresponds to a cost of 1.9 INR per kWh of avoided unmet demand. Intermediate levels of conservatism are more cost effective, reducing to 1 INR per kWh of avoided unmet demand at 50% conservatism, albeit with far more unmitigated unmet demand.



(a) Installed capacity.



(b) Annual electrical energy production.

ECh B-CCHP StoreE StoreT DG PV GridE

Figure 5.26 Comparison of mean model a) installed capacity and b) annual electricity production to results given by various models aimed at improving power interruption resilience.

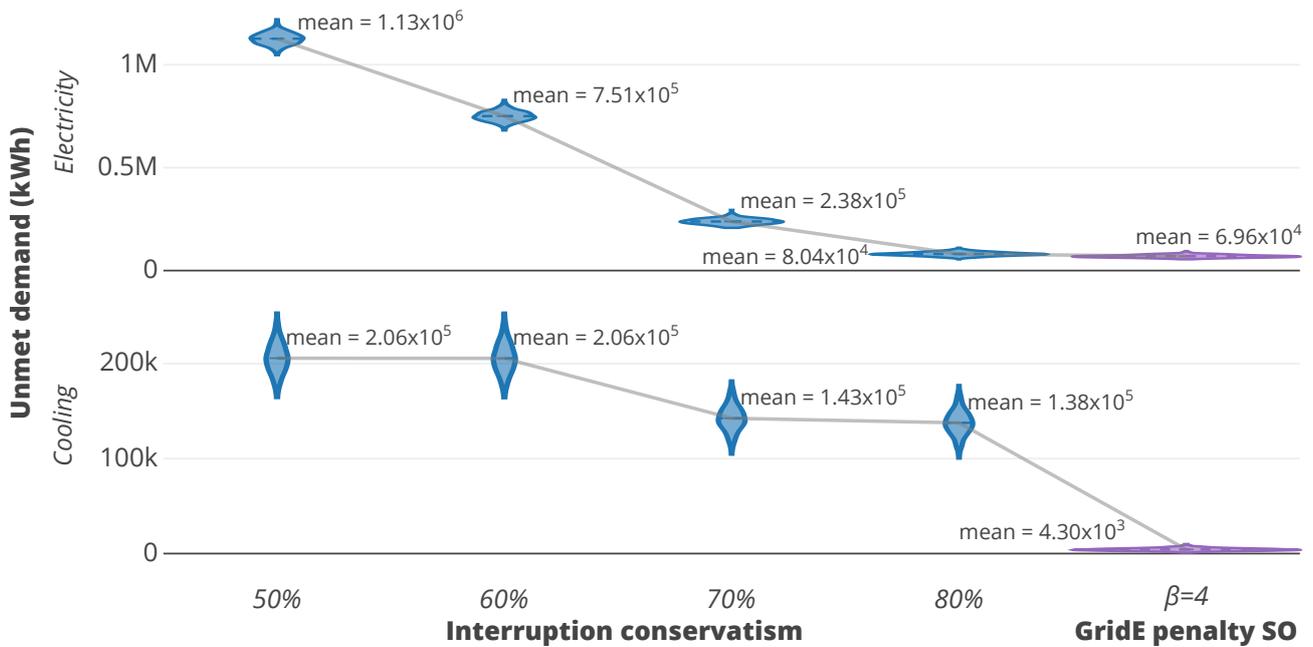


Figure 5.27 System unmet cooling and electricity demand distributions when running OOS optimisation tests, including the introduction of electricity power supply interruptions, on Bangalore case study models which have incorporated power interruption resilience.

Table 5.7 Investment cost ( $\times 10^7$  INR) for optimal solutions derived from Bangalore models aimed at improving power interruption resilience.

Mean	Interruption conservatism				GridE penalty SO		
	50%	60%	70%	80%	$\beta = 0$	$\beta = 0$	$\beta = 4$
5.01	6.18	6.29	6.76	7.27	4.92	6.32	7.34

## 5.6 Discussion

Although there is clear variation in demand exhibited in the generated scenarios, the impact on technology choice is small. This is likely to be caused by the benefit of conventional technologies when demand uncertainty is inherent in a system. Renewable technologies are non-dispatchable and district systems require management of multi-energy distribution; both limit flexibility. Conventional technologies are also cheaper. It is thus not surprising that risk-unaware investment decisions are similar to risk-aware ones. The proposed investments are not identical, however. In Bangalore, there is a reduction in PV capacity when accounting for demand uncertainty. In Cambridge there is an increase in energy storage capacity. Both indicate the clear need for increased flexibility in a district system, but are dwarfed by the additional capacity deemed necessary in conventional technologies. Indeed, much of the system flexibility is achieved by investment in larger technologies which have a low capacity factor in any realised demand scenario.

A particular difference between the two case studies is the cost incurred in accounting for demand uncertainty. In both, the total system cost increases. In Cambridge this is led by an increase in investment cost, allowing operation costs to reduce relative to the 500 input

scenarios. In Bangalore, the investment cost decreases relative to the reference model, but the operation costs increase considerably as a result. Whether one is preferable to another depends on the decision maker, who may be responsible for only meeting one of the cost streams.

Introducing risk aversion increases the **OFV** mean and decreases the standard deviation in both studies. This is the expected impact from the introduction of the **CVaR** risk measure in the objective function. But, it does not lead to perceptible changes in investment decisions in the Bangalore case. Instead, the operation schedule is changed to mitigate the worst case costs. This may not seem useful at first, but it does indicate that the investment decision is the most robust choice available for the system. The small change in technology capacities in the Cambridge case indicates that the flexibility introduced by heat storage is not as useful as the increase in conventional technology capacity.

Although there are differences in results between **SO** and single scenario optimisation, the impact of this on system robustness can only be understood through **OOS** tests. As such, the omission of this step in many existing studies is a clear shortcoming. The impact is positive in both case studies, whether compared to the reference model or the worst case single-scenario demand/**OFV** models. The quantity of unmet demand is noticeable in the reference model, constituting 5 – 10% of the annual demand, depending on the case study and energy type. This is reduced to less than 1% by the introduction of **SO**, but could be further reduced to negligible levels if the **SO** model were built over the full time horizon. Unfortunately, such models would be intractable, so the possible benefits are not known in this thesis.

Unmet demand is not only affected by uncertainty in demand, but also uncertainty in the availability of supply. In fact, power interruptions, which are an almost daily occurrence in Bangalore, have a much greater impact on systems which do not expect them than the impact of uncertain demand. As with demand uncertainty, it is possible to mitigate the effect of power interruptions. Doing so leads to a shift in technology choice, but this choice is still conventional, primarily towards building-level **DGs**. This result is reached whether accounting for power interruptions explicitly, with an increasingly conservative perspective on supply availability in a single-scenario model, or by reducing reliance on the grid by updating the three-step methodology to include electricity purchases as a risk in **SO**.

Unlike the models which only consider demand uncertainty, handling power interruptions comes at a considerable increase in technology investment costs. However, de-risking power interruptions is evident in technology choices made for commercial properties in Bangalore, whereby **DGs** meet **GridE** supply shortfall. Indeed, the building from which demand data was acquired has two diesel generators to handle power interruptions. This additional investment could be justified if considering it as a cost per unit of mitigated unmet demand, but this puts pressure on the optimisation results to be realistic. Instead, the magnitude of unmet demand is likely to be much lower in all models, as evidenced by the impact of considering the full timeseries. Nonetheless, the results show that technology choice and sizing varies depending on the degree to which uncertainty is internalised in a model, leaving it to a decision maker to decide the risks they are, and are not, willing to take.

## 5.7 Conclusions

Decision making under uncertainty can be improved by the combination of new sampling and optimisation techniques. This chapter has proposed a three-step method by which district energy systems can be designed to be more robust to demand uncertainty than traditional, single scenario models. These three steps are **scenario generation (SG)**, **scenario reduction (SR)**, and **scenario optimisation (SO)**. In **SG**, a multivariate nonparametric sampling method is used to produce 500 future demand scenarios from historic building-level stochastic consumption data. Using **SR**, a scenario subset is chosen to ensure a tractable **SO**, without misrepresenting the probability distribution of our 500 initial scenarios.

These three steps have been applied to illustrative case studies in Cambridge, UK, and Bangalore, India. By using **out-of-sample (OOS)** scenarios, unmet demand has been quantified for **SO** and single scenario technology investment portfolios. Robustness to unmet demand increases by two orders of magnitude in both cases when using **SO**-derived technology capacities over those derived from a scenario representing business-as-usual optimisation. In the Cambridge case, this robustness comes at a high cost of  $+2 \times 10^5$  **GBP** investment. In Bangalore, greater robustness can be achieved at a lower investment cost than given by the mean scenario, but given the expectation that load shedding will be required if a low probability scenario is realised. By direct application to a masterplan-level site (Cambridge), this method has been demonstrated as a viable tool for use by decision makers to improve decision robustness.

The Bangalore system was also subjected to power interruptions, which is an existing issue for buildings connected to the national electricity grid in the city. The resilience of investment portfolios was first analysed, which constitutes those which were completely unaware of the possibility of interruptions. The increase in unmet demand suggests that the investment portfolios are not resilient. Consequently, the risk being mitigated in **SO** was updated to reduce dependence on grid electricity, such that the resulting system maintains low levels of unmet demand in **OOS** tests even when power interruptions are included.

Contrary to existing literature, district energy systems are not considered cost-optimal in this chapter. Conventional energy technologies are preferred, such as building-level **natural gas boilers (NGBs)**, **electric chillers (EChs)** and, when power interruptions are introduced, **diesel generators (DGs)**. There is some reliance on rooftop **solar photovoltaic panel (PV)** in Bangalore. However, increased robustness leads to reduction in capacity. If district systems are to be included, due to perceived additional benefits or regulatory requirements, then the objective must change. Nevertheless, it can be readily seen that doing so is unlikely to aid in the endeavour to improve system robustness to unmet demand.



# Chapter 6

## Whose objective is it anyway?

In district energy system optimisation, the aim is often to minimise cost for a given set of demand. Depending on the stage of development, this may refer to the cost of operating an existing system (Ren *et al.*, 2010), of investing in a system for a new development (Keirstead *et al.*, 2010), or of upgrading existing infrastructure (Farzaneh *et al.*, 2016; Omu *et al.*, 2015). In all such cases, the objective function of the optimisation model is a monetary cost function of technology capacity and/or dispatch.

At a district level, the decision being made is between energy conversion within buildings (decentralised) or in a district-wide energy centre (centralised). Centralised district energy systems are considered an improvement on incumbent designs due to their energy efficiency and the possibility of emissions reduction. Studies have strengthened this claim, reporting 23% - 50% reductions in carbon emissions as a result of meeting district energy demand with a centralised system (Li *et al.*, 2016a; Morvaj *et al.*, 2016; Omu *et al.*, 2013). This possibility of emissions reduction has led to policies incentivising centralised systems, such as the 2025 target of at least 25% of thermal demand to be met by centralised systems by 2025 in London, UK (GLA, 2016, Ch. 5).

However, the economic viability of centralised systems is a point of contention. Some studies conclude with considerable possible savings (Jennings *et al.*, 2014; Mehleri *et al.*, 2012), while others only realise minor savings (Buoro *et al.*, 2014) or even an increase in system cost (Haikarainen *et al.*, 2014; Omu *et al.*, 2015). The results of Chapter 5 agree with these latter studies: a monetary objective is insufficient to incentivise the implementation of centralised district energy systems. However, if district energy systems are optimized only to minimize monetary cost, carbon emissions could more than double (Jennings *et al.*, 2014).

To balance monetary cost and the increasingly important carbon emissions, studies often implement a linearised multi-objective optimisation. To do so, carbon emissions are either monetised and included in the monetary cost objective, or the total allowable carbon emissions is progressively limited whilst maintaining a purely monetary objective. Monetised emissions may directly represent a 'carbon tax' (Mehleri *et al.*, 2012) for CO<sub>2</sub> emissions, or a perceived cost to society resulting from the effects of air pollutants, such as their negative health impacts (Chen *et al.*, 2016; Omu *et al.*, 2015; Zhang *et al.*, 2015a). Whether a tax or a social cost, the recommended value to assign emissions can vary quite significantly (Hourcade

*et al.*, 2017; IWG, 2016; BEIS, 2018b). Instead, emissions can be progressively limited whilst minimising only monetary cost, also known as the ' $\epsilon$ -constraint method', allowing a modeller to examine the trade-off between the two on a cost-carbon Pareto front (Majewski *et al.*, 2017; Morvaj *et al.*, 2016; Sharafi and ElMekkawy, 2014; Wang *et al.*, 2017; Wei *et al.*, 2016; Zangeneh *et al.*, 2009). Although a subjective choice must then be made on which solution along the Pareto front is favoured, the  $\epsilon$ -constraint method provides a much clearer understanding of the monetary impact of any marginal reduction in carbon emissions.

In addition to monetary cost and carbon emissions, there is a degree of economic risk, regulatory uncertainty, and technology lock-in associated with centralising energy production at district scales (Hawkey *et al.*, 2013; Kelly and Pollitt, 2010). On a technological level, a more centralised district system requires the introduction of more interdependent systems. With this interdependence comes the risk of a system which is less resilient (Woods, 2015). In other words, a centralised district energy system can be more vulnerable to sudden perturbations or future evolutions of demand due to the effect of cascading failures along interlinked energy systems.

The literature generally supports this hypothesis: more conventional technologies provide greater resilience to future variations in demand (Akbari *et al.*, 2016; Majewski *et al.*, 2017). Indeed, whether resilience is defined as robustness or graceful extensibility, results from Chapter 5 show that the cost-optimal solution is further weighted towards conventional decentralised technologies when accounting for demand uncertainty or possible **national grid electricity (GridE)** interruptions.

The focus of a design will depend on the actors involved in the decision-making process, with inevitable compromise resulting when objectives compete. However, there is only a limited understanding of this compromise in the literature; studies generally only consider one objective, or at most two. Multi-objective optimisation usually focusses on either cost and carbon (e.g. Li *et al.*, 2016a; Morvaj *et al.*, 2016; Zhang *et al.*, 2015b) or cost and resilience (e.g. Bucciarelli *et al.*, 2018; Mavromatidis *et al.*, 2018c; Tanaka *et al.*, 2017; Zhou *et al.*, 2018).

When monetary cost, carbon emissions, and risk aversion are part of the objective in district energy systems, the impact of the compromise is never fully explored. For instance, Pazouki and Haghifam (2016) considered the marginal impact on the system design when varying the importance of resilience, but ignored any variation in the importance of carbon emissions. Conversely, Mavromatidis *et al.* (2018a) varied a limit on carbon emissions in a cost minimising **scenario optimisation (SO)** model, but did not consider different types of uncertainty nor degree of risk aversion. Majewski *et al.* (2017) has made the most progress in considering the impact of competing objectives, by considering the additional impact of compounding uncertainties on design whilst also varying the relative importance of carbon emission and monetary cost reduction using **robust optimisation (RO)**. However, they made no attempt to validate the robustness of their design using **out-of-sample (OOS)**, such that the benefits in the direction of a resilience objective are unknown.

To understand the sensitivity in investment decisions to the addition of different decision makers, it is thus necessary to test individual case studies against a number of objectives and, where uncertainty is involved, to validate designs against **OOS** scenarios. In this chapter,

such a process is undertaken, focussing on eight potential decision makers in the process of designing a district energy system:

1. The cost-focussed client, who wishes to see the balance of investment and operation costs minimised;
2. The cost-focussed developer, whose objective is to avoid higher investment costs, even if it is at the expense of higher operation costs;
3. The cost-focussed occupant, whose objective is to avoid higher operation costs, without any power over the district investment decisions;
4. The carbon-focussed client, who is willing to pay a cost premium to reduce carbon emissions resulting from the system operation;
5. The cost-focussed, carbon-aware client, who is obliged to either limit carbon emissions associated with their development or pay for any emissions they emit;
6. The carbon-focussed, cost-aware developer, who is obliged to reduce carbon emissions resulting from the system operation, but has a limit on capital investment which can be afforded;
7. The cost-focussed, district-constrained client, who has to provide a centralised system to meet a proportion of a district's thermal demand;
8. The risk-aware client, who wants to ensure that the risk of realising high cost, carbon, or unmet demand futures is kept low.

Although the list is far from exhaustive, the choice of these eight decision makers aims to fully encompass the cost-carbon-resilience decision-making nexus as far as is possible in district energy design. The objective of each decision maker is applied to the two case studies found in Chapter 5. Reformulations of the objective function, and the resulting impact on case study model results, are detailed in Section 6.1. In Section 6.2, OOS tests are run to analyse system robustness to demand uncertainty, to validate the designs resulting from risk-awareness and to understand the unintended impact of single scenario objectives.

## 6.1 Reformulating the objective function

Although typical in district energy system research, there are many other forms of the objective function than that given in Equation 3.26. The objective function is the weighted sum of various decision variables. On updating the objective function, the model decision variables remain the same, only their contribution to the objective function differs. Equally, decision-maker objectives may be influenced by the introduction of new constraints, to ensure that policy, regulation, or site specific objectives can be incorporated.

A typical district energy system optimisation model aims to minimise system annualised monetary cost, both from initial capital investment and subsequent technology operation.

Each of the eight decision makers defined in the previous section may require a reformulation of the objective function, beyond that which is typically used in [mixed integer linear programming \(MILP\)](#). In this section, each distinct formulation is defined, alongside any additional constraints which are required. Parameter, set, and decision variable terminology is given in Table 6.1.

Table 6.1 Decision variables, sets, and parameters defined in this chapter.

## (a) Decision variables.

$\hat{\mathbf{E}}$	Energy capacity	$\mathbf{cost}^{op}$	Operational costs
$\mathbf{E}^+$	Energy carrier production	$\mathbf{cost}^{inv}$	Investment costs
$\mathbf{E}^-$	Energy carrier consumption	$\mathbf{P}$	Technology purchase switch
$\mathbf{cost}$	System cost	$\mathbf{slack}^+$	Unmet demand slack
$\hat{\mathbf{R}}^{area}$	Resource collection area	$\mathbf{slack}^-$	Unused supply slack
$\hat{\mathbf{S}}$	Stored energy capacity	$\zeta$	Value at risk (VaR)
$\mathbf{E}^{ex}$	Energy carrier export	$\eta$	Auxiliary risk variable

## (b) Sets.

$n \in nodes$	Geographic nodes in the system
$x \in techs$	Available technologies
$t \in timesteps$	Operational timesteps
$c \in carriers$	Energy carriers
$k \in costs$	Cost classes, where $k_1$ = monetary and $k_2$ = carbon emissions
$s' \in S'$	Reduced scenarios
$y \in Y$	Capacity decision variables $\{\hat{\mathbf{E}}, \hat{\mathbf{R}}^{area}, \hat{\mathbf{S}}, \mathbf{P}\}$

## (c) Parameters.

$bigM$	Large numeric value associated with use of slack variables
$cost^{ex}$	Cost associated with exporting a carrier
$cost^{prod}$	Cost associated with producing a carrier
$cost^{con}$	Cost associated with consuming a carrier
$cost^y$	Capacity cost associated with capacity decision variable $y$
$TSF$	Timeseries scaling factor
$W$	Scenario weight
$\alpha$	Confidence level
$\beta$	Risk aversion level

### 6.1.1 The Pareto front

Existing studies which do not consider a cost-focussed client often present a range of optimal solutions. In non-deterministic modelling, these solutions exist within a convex hull, for which there is a 'Pareto front' describing the dominant solutions (Figure 6.1) (Alarcon-Rodriguez *et al.*, 2010). Within the hull there will be non-dominant solutions for which there is another (dominant) solution which achieves a reduced value in the x/y axis, for the same value in the y/x axis. In MILP, dominant solutions will always be achieved, so only the Pareto front will be defined. Most studies find that a cost-carbon Pareto front follows a trend similar to that illustrated in Figure 6.1 (Majewski *et al.*, 2017; Morvaj *et al.*, 2016; Sharafi and ElMekkawy, 2014; Wang *et al.*, 2017; Wei *et al.*, 2016; Zangeneh *et al.*, 2009). The decision maker must then make a subjective trade-off to assign preference to one of the dominant solutions.

In this section, the Pareto front of dominant solutions is given for a number of competing objectives. The three objectives of interest are: minimising carbon emissions, minimising investment cost, and minimising operation cost. The two-dimensional Pareto front has therefore been extended in the following sections to include the values corresponding to a third objective, shown on the solution markers as colours varying along a colour scale.

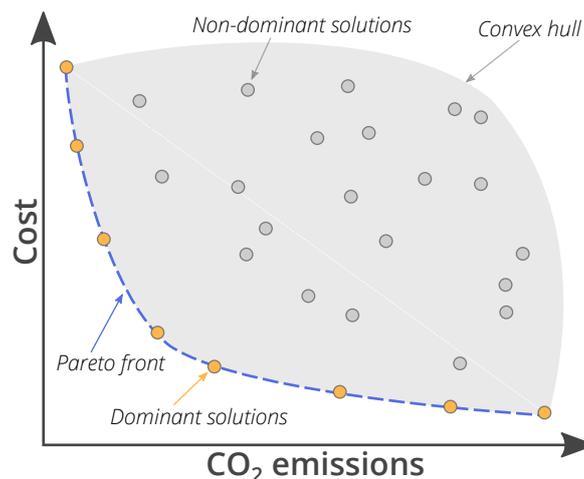


Figure 6.1 Representation of a convex hull, containing non-dominant solutions resulting from multi-objective optimisation, and a Pareto front on which the dominant solutions lie. The dashed line is described by the dominant solutions.

### 6.1.2 The cost-focussed client

The objective function for monetary cost minimisation (Equation 6.1) is used most often in energy system optimisation. This may require the consideration of investment and operation costs ( $\phi = 1$ ) (e.g. Gabrielli *et al.*, 2018; Morvaj *et al.*, 2016; Voll *et al.*, 2015), or only operation ( $\phi = 0$ ) (e.g. Orehounig *et al.*, 2015; Wang *et al.*, 2015). In the latter case, technology capacities are pre-defined, either to emulate an existing system or to be varied by user-defined scenarios. When investment costs are included it is necessary to scale them to the operating period being modelled. If modelling a full year, the investment cost is annualised. The annualised cost

accounts for both the technology lifetime  $L$  and the financial interest rate  $R$  by applying the annualisation factor  $\frac{R}{1-(1+R)^{-L}}$  to the investment cost. This annualisation factor can be further generalised to a **timeseries scaling factor (TSF)**  $\frac{T}{8760} \times \frac{R}{1-(1+R)^{-L}}$ , where  $T$  is the number of operating hours in the model under consideration. Using the **TSF**, the investment and operation costs will be weighted equally in the objective function, irrespective of the number of hours represented in the model.

$$\begin{aligned}
\min \quad & \phi \times TSF \times \sum_{n,x,y} \left( \mathbf{y}_{n,x} \times cost_{n,x,k_1}^y \right) \\
& + \sum_{n,x,c,t} \left( \mathbf{E}_{n,x,c,k_1,t}^+ \times cost_{n,x,k_1,t}^{prod} - \mathbf{E}_{n,x,c,k_1,t}^- \times cost_{n,x,k_1,t}^{con} + \mathbf{E}_{n,x,c,k_1,t}^{ex} \times cost_{n,x,k_1,t}^{ex} \right) \quad (6.1) \\
& + bigM \sum_{n,c,t} \left( \mathbf{slack}_{n,c,t}^+ - \mathbf{slack}_{n,c,t}^- \right)
\end{aligned}$$

The results from the cost-focussed client model are considered as the ‘baseline’ cost model. For both the Bangalore and Cambridge case studies, the baseline model was optimised in Chapter 5, in which it was referred to as the ‘mean’ model. There is little monetary incentive to incorporate centralised systems into the baseline models, although **solar photovoltaic panels (PVs)** are installed on up to one third of the available rooftop area in the Bangalore district.

### 6.1.3 The cost-focussed developer or occupant

#### Objective function formulation

When an unequal weighting is desired in the objective function, it is possible to update  $\phi$  to vary in the range  $[0, \infty)$ . A higher  $\phi$  will disproportionately weight investment cost, such that the objective function value will include reduced investment costs at the expense of high operation costs. However,  $\phi$  is an abstract parameter whose scale cannot readily be understood when applied to the objective function.

Instead of subjectively assigning values to  $\phi$ , the investment cost can be removed from the objective function and limited within a constraint referred to as the  $\epsilon$  constraint (Equation 6.2) (Farzaneh *et al.*, 2016; Mavromatidis *et al.*, 2018c). Taking the baseline optimal investment cost as that which is achieved when minimising Equation 6.1, the maximum allowed investment cost is limited to the baseline cost multiplied by the factor  $\epsilon$ . A low  $\epsilon$  will reduce investment costs at the expense of increased operation costs, and vice versa for a high  $\epsilon$ . This may seem more practical to decision makers, as developments will have budgets that must be met. By varying the magnitude of  $\epsilon$ , a different optimal configuration of decision variables

will be reached. However, the overall cost (investment + operation) will be greater if the configuration does not match the result obtained by the cost-focussed client.

$$\begin{aligned}
 \min \sum_{n,x,c,t} & \left( \mathbf{E}_{n,x,c,k_1,t}^+ \times \text{cost}_{n,x,k_1,t}^{prod} - \mathbf{E}_{n,x,c,k_1,t}^- \times \text{cost}_{n,x,k_1,t}^{con} + \mathbf{E}_{n,x,c,k_1,t}^{ex} \times \text{cost}_{n,x,k_1,t}^{ex} \right) \\
 & + \text{big}M \sum_{n,c,t} (\text{slack}_{n,c,t}^+ - \text{slack}_{n,c,t}^-) \\
 \text{s.t.} \sum_{n,x,y} & \mathbf{y}_{n,x} \times \text{cost}_{n,x}^y \times \text{TSF} \leq \epsilon \times \text{cost}_{k_1}^{inv,baseline}
 \end{aligned} \tag{6.2}$$

### Application of the objective function

By applying two extreme decision-maker priorities to the case study models (respectively prioritising investment and operation cost), it is possible to calculate the limits on system cost. Figure 6.2 details the investment and operation costs associated with these two extreme models. Between the developer ( $\min(\text{cost}^{inv})$ ) and the occupant ( $\min(\text{cost}^{op})$ ), the Bangalore case study can vary in investment cost by two orders of magnitude while the Cambridge case study can vary by five orders of magnitude. The Cambridge baseline model technology configuration matches the cost-focussed developer results, leading to the same **objective function value (OFV)**, whereas the lowest investment cost in the Bangalore case study is a third less than that given by the baseline model. The magnitude of operation costs is much larger than for investment costs, so although the percentage variation is relatively low, the magnitude of variation is still significant. Indeed, the occupant could reduce their annual expenditure by  $6 \times 10^6$  GBP and  $8 \times 10^7$  INR in the Cambridge and Bangalore districts, respectively.

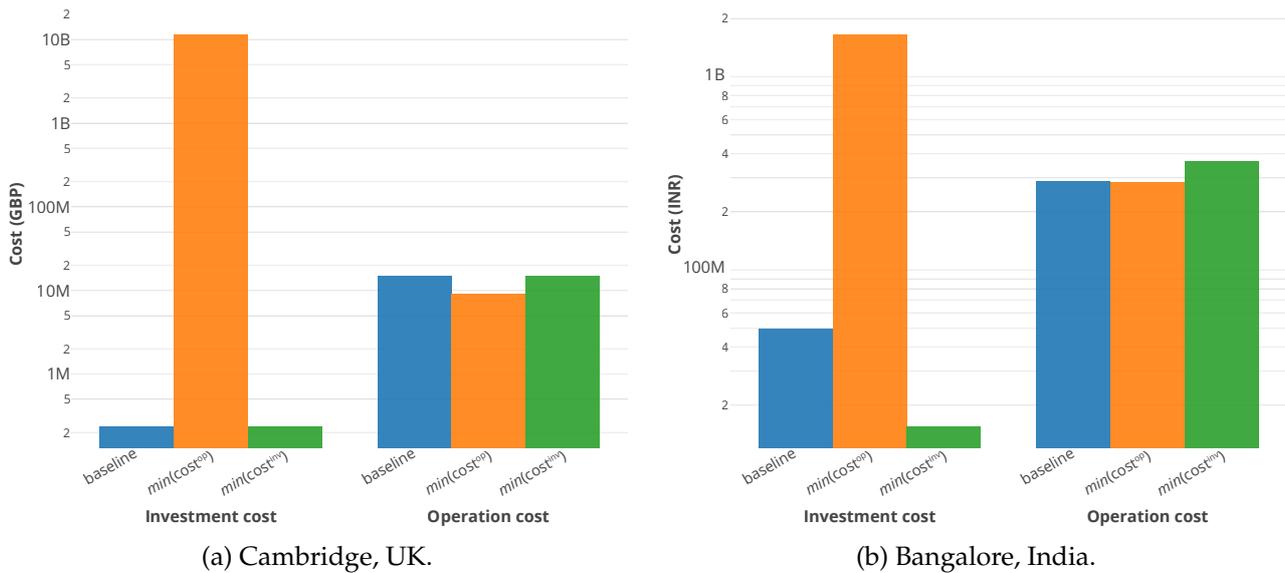


Figure 6.2 Comparison of optimal investment and operation costs for Cambridge and Bangalore case studies when balancing operation and investment costs (baseline), minimising only investment ( $inv$ ) cost, and minimising only operation ( $op$ ) cost in the objective function. A base-10 logarithmic scale is used for the  $y$  axis.

$\epsilon = 1$  is given for the optimised investment cost of the baseline models.  $\epsilon$  values of 1 and 0.3 are associated with the Cambridge and Bangalore cost-focussed developer models, respectively, given the difference in optimised investment cost relative to the baseline models shown in Figure 6.2. Models were run using  $\epsilon$  values starting at these minimum values, increasing by 0.1 until an  $\epsilon$  of two was reached, i.e. twice the baseline model investment cost. The Bangalore Pareto front in Figure 6.3b shows a distinct discontinuity at the baseline model investment cost. This is an expected shape for the Pareto front, as the operation and investment costs would form a convex hull which would inevitably include a vertex at which the optimum, balanced system cost is found. As the Cambridge baseline model has already minimised investment cost, the same discontinuity does not exist in Figure 6.3a. However, the linear variation of the Pareto front from  $\epsilon = 1$  to 2 is consistent with the same  $\epsilon$  range in the Bangalore model.

The technology portfolio varies with  $\epsilon$ . This is particularly the case in the Bangalore model, in which there are different choices made either side of the baseline model results (Figure 6.4b). With a more constrained investment cost, PV capacity decreases, until reaching the minimum investment cost model where no PV is considered. As the value of  $\epsilon$  in Bangalore increases above one, the electricity technology investment decision remains static. Instead, a centralised cooling system is incrementally introduced. This involves increasing the capacity of an energy centre ECh (EC-ECh). Although the capacity of the EC-ECh only reaches a quarter of the cumulative size of the building-level electric chillers (EChs), its greater efficiency means it can meet more than two thirds of cooling demand in the  $\epsilon = 2$  model. There is far less variation in the Cambridge technology portfolio seen in Figure 6.4a, which introduces some PV capacity as  $\epsilon$  increases. However, even at  $\epsilon = 2$ , there is little noticeable change. This is expected, when considering the large increase in investment cost required to reach the cost-focussed occupant result.

As the investment cost reduces, and the operation cost reciprocally increases, the carbon emissions increase. The increase in emissions is in proportion to the increase in operation cost. Namely, the decrease in emissions is caused by the investment in PVs. This can be seen in Bangalore, where there is no change in emissions between  $\epsilon = 1$  and 2, even though there is a reduction in use of the less efficient EChs in favour of the EC-ECh.

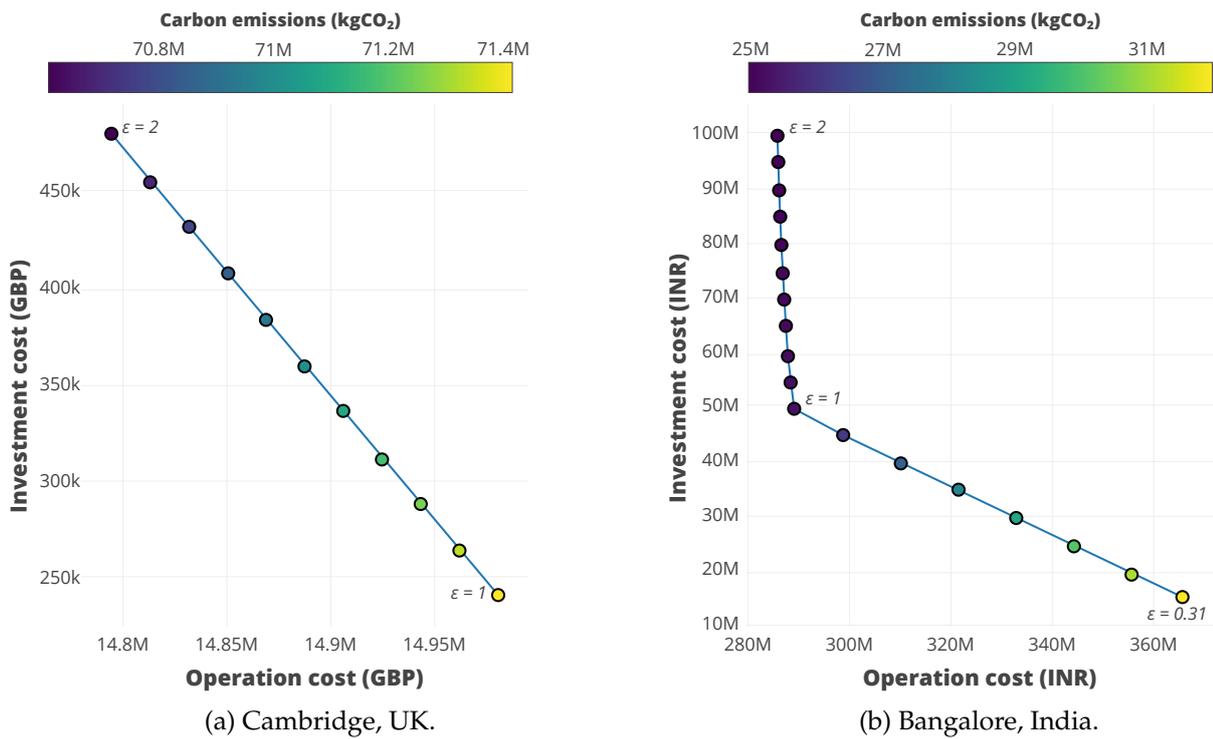


Figure 6.3 Investment-operation cost Pareto front following monetary cost minimisation of both case studies with  $\epsilon$ -constrained investment cost. Carbon emissions associated with each model run are given by the marker colour.

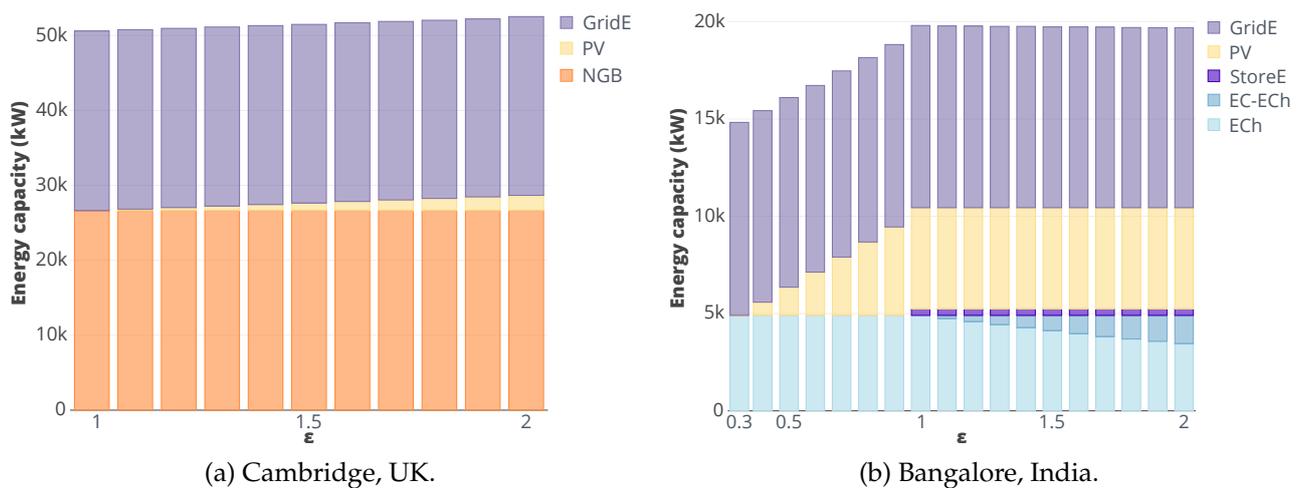


Figure 6.4 Installed energy capacity following monetary cost minimisation of both case studies with  $\epsilon$ -constrained investment cost.

### 6.1.4 The carbon-focussed occupant

On minimising carbon emissions instead of monetary cost, the objective function becomes Equation 6.3. Only timeseries emissions are included in this objective function, hence the occupant being the decision maker in this instance. Embodied emissions are not included. Time-varying carbon emissions are applied to system resources: natural gas, diesel, biomass, and grid electricity. As such, investment and operation of renewable energy technologies and storage technologies has no impact on the OFV.

$$\min \sum_{n,x,t} \left( \mathbf{E}_{n,x,t}^- \times \text{cost}_{n,x,k_2,t}^{\text{con}} \right) \quad (6.3)$$

The results from the carbon-focussed client model are considered as the ‘baseline’ carbon model. Figure 6.5 shows that, for both case studies, a centralised system becomes more viable in the baseline carbon model, when compared to the baseline cost model. Complete independence from the electricity grid is achieved in Bangalore, in favour of biomass and diesel fuel use. Although **diesel generators (DGs)** are often considered as the highly polluting back-up electricity source, in Bangalore they are able to produce electricity with a lower emission factor ( $0.63\text{kgCO}_2/\text{kWh}_e$ ) to the national grid ( $0.7\text{kgCO}_2/\text{kWh}_e$ ) (cBalance, 2012). Such a high emissions factor for **GridE** (approximately twice that of the UK, for example) is a function of both the Karnataka state fuel mix and its high transmission losses ( $\sim 13\%$ ) (BESCOM, 2017). Figure 6.6 shows that the **combined cooling, heat and power plant (CCHP)** is thermally led for half the **typical days (TDs)**. Otherwise, it makes heavy use of **cold water StoreT (StoreT-C)** to maximise electrical output. Additional electrical requirements are met by building-level **DGs**. Little use is made of **electrical battery storage (StoreE)**, usually charging at peak **PV** output and discharging at peak electricity demand, since there is a 2-3 hr lag between the two peaks.

Of note is the minor use of an **ECh** in TD 12. In this **TD**, the **PVs** overproduce electricity. Since there is no investment cost in this objective function, there is no disadvantage for the model to invest in an additional technology to make best use of this excess, namely to meet electricity demand by **ECh** instead of by using the **biomass fuelled CCHP (B-CCHP)**, which does have some carbon emissions associated with its use.

In Cambridge, Figure 6.7 shows that the **combined heat and power plant (CHP)** is only used when **PV** output is low, constituting three of the **TDs**. This also corresponds to high heat demand, allowing the **CHP** to follow heat demand whilst also counteracting low **PV** output. The **solar thermal panel (ST)** is installed, but is used only rarely to meet heat demand at the same time as heat is produced. Indeed, **ST** output peaks at the same time heat demand is particularly low. Instead, the **hot water StoreT (StoreT-H)** is utilised. However, much of the heat stored in the summer is dissipated from the heat store before the winter periods, due to standing losses of the vessel. The **ground source heat pump (GSHP)** is used as a baseload heat supply technology. Although *quality* has not been considered in these models, the relatively low temperature output of a **GSHP** would support its use as a baseload technology in a real system.

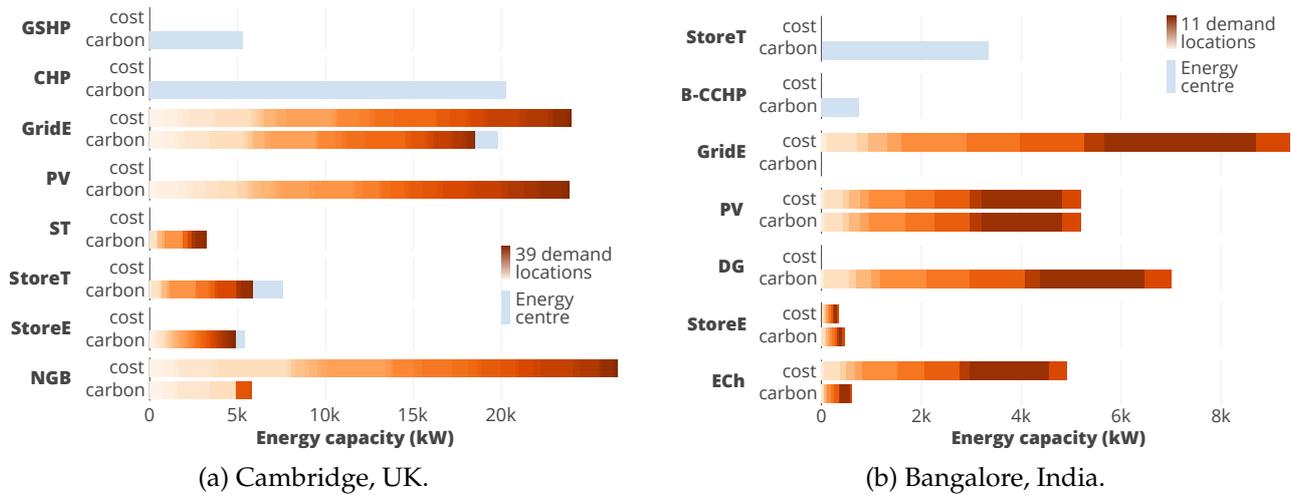
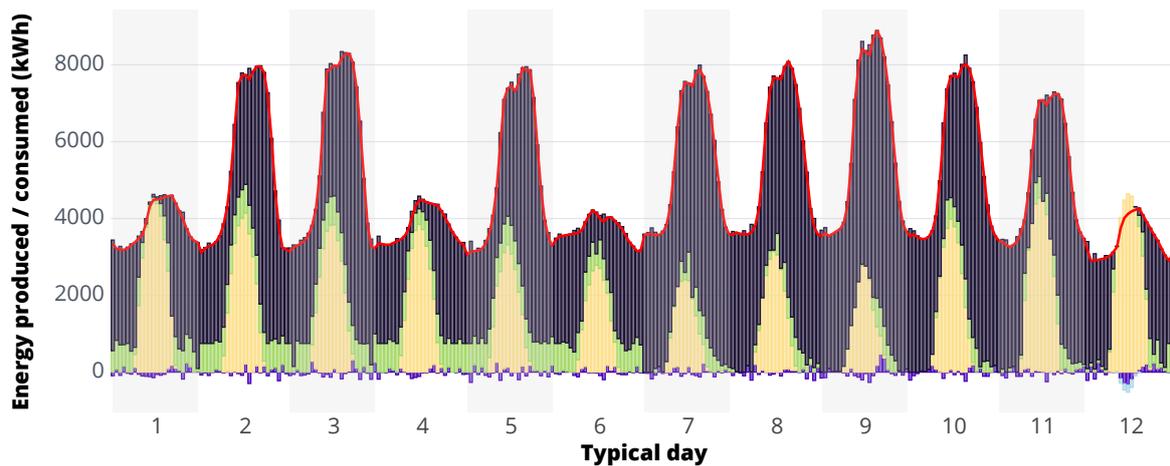
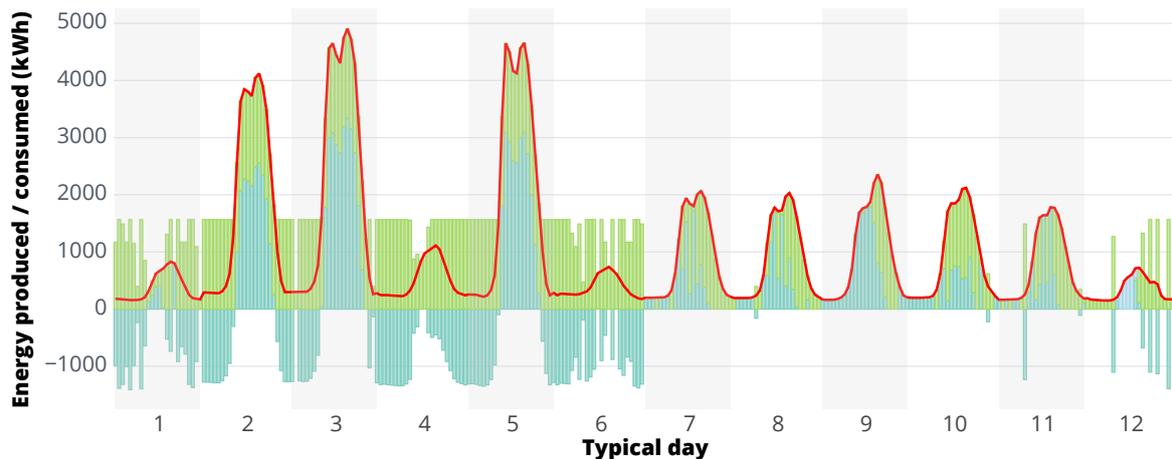


Figure 6.5 Installed capacity of technologies to achieve the optimal OFV in both **cost** minimisation and **carbon** emission minimisation models. Building-level technologies have been aggregated over all demand nodes. The contribution from each demand node to the total technology capacity has been differentiated with a colour gradient. Although, in some cases (e.g. **natural gas boiler (NGB)** in (a), carbon minimisation), not all nodes have installed capacity.



(a) Electricity.

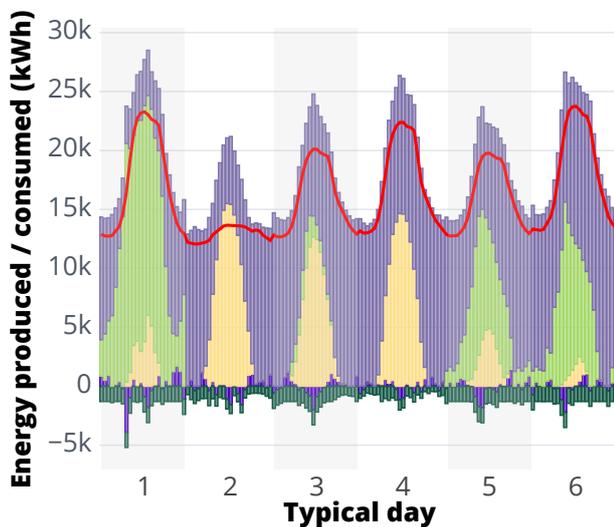
Figure 6.6 Timeseries energy production and consumption in the **Bangalore** case study for technologies operating in each typical day, as given by the optimal model solution following carbon emission minimisation.



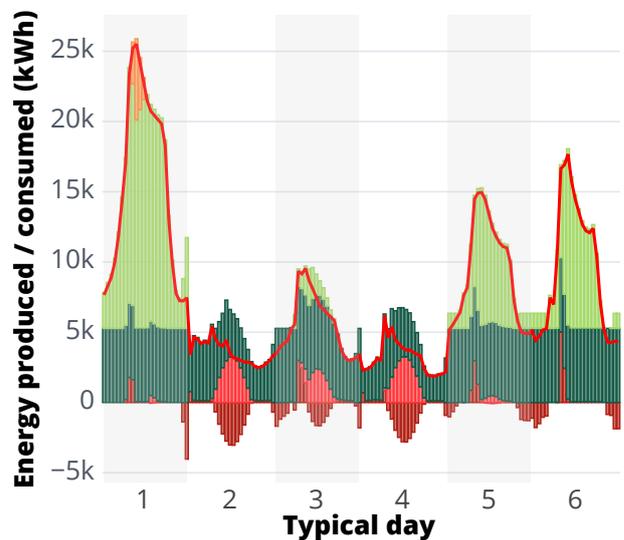
(b) Cooling.

Legend: StoreE, StoreT, PV, B-CCHP, ECh, DG, GridE, demand

Figure 6.6 (cont.) Timeseries energy production and consumption in the **Bangalore** case study for technologies operating in each typical day, as given by the optimal model solution following carbon emission minimisation.



(a) Electricity.



(b) Heat.

Legend: StoreE, StoreT, ST, PV, GSHP, CHP, NGB, GridE, demand

Figure 6.7 Timeseries energy production and consumption in the **Cambridge** case study for technologies operating in each typical day, as given by the optimal model solution following carbon emission minimisation.

### 6.1.5 The cost-focussed, carbon-aware client

**MILP** studies rarely model a carbon-focussed occupant. Instead, carbon emissions are incorporated into the model either 1) by addition into a monetary cost objective by using a ‘cost of carbon’ (Equation 6.4) (Morvaj *et al.*, 2016), or 2) by constraining the total allowed emissions for the system under consideration (Mavromatidis *et al.*, 2018a).

#### Objective function formulation

**Cost of carbon** Two methods for assigning cost can be considered: the market value of emissions, as given by financial incentives such as an emissions trading scheme, and the social cost of carbon (SCC). Electricity generators in Europe are obliged to participate in the emissions trading scheme (EU-ETS). The EU-ETS is a mechanism aimed at creating a market value for reduced emissions from electricity production. However, its poor performance in recent years has led to the introduction of a ‘price floor’ in the UK, which increases the market value of carbon by 18 GBP/tCO<sub>2</sub> up to 2020 (Hirst, 2018).

Emissions trading schemes create a market for carbon monetisation and, like any commodity, values can fluctuate. Additionally, no emissions trading scheme exists yet in India. The same cannot be said for the SCC, which is instead the perceived cost to society due to the emission of CO<sub>2</sub> (or other gases with global warming potential) into the atmosphere (IWG, 2016). The SCC considers the financial losses incurred at country level by the direct effects of climate change on health, forests, and water systems (Akbar *et al.*, 2014; IWG, 2016). The IWG (2016) recommend an SCC of approximately 14 GBP/tCO<sub>2</sub>, which was used by Omu *et al.* (2015) to understand the cost of district-scale emitted CO<sub>2</sub>. This closely matches the value used by Mehleri *et al.* (2012) of 17€/tCO<sub>2</sub>, but recommendations differ on the value to use. In fact, the uncertainty with which the SCC is created and the unlikelihood of higher values ever filtering through to decision making may make it an unworkable metric (Taylor, 2017). In India, although the recommended cost of carbon is 40 USD/tCO<sub>2</sub> today, a more politically viable cost is expected to be closer to 5 USD/tCO<sub>2</sub> (Hourcade *et al.*, 2017). To cover this uncertainty and encompass the market and social value of carbon emission reduction, varying costs of carbon can be tested within Equation 6.4. A cost of carbon range, covering both the current/expected SCC and market values, is given in Table 6.2.

$$\begin{aligned}
 \min \quad & \phi \times TSF \times \sum_{n,x,y} \left( \mathbf{y}_{n,x} \times \text{cost}_{n,x,k_1}^y \right) \\
 & + \sum_{n,x,c,t} \left( \mathbf{E}_{n,x,c,k_1,t}^+ \times \text{cost}_{n,x,k_1,t}^{prod} + \mathbf{E}_{n,x,c,k_1,t}^{ex} \times \text{cost}_{n,x,k_1,t}^{ex} \right. \\
 & \left. - \mathbf{E}_{n,x,c,k_1,t}^- \times (\text{cost}_{n,x,k_1,t}^{con} + \text{cost}_{n,x,k_2,t}^{con} \times \text{cost}_{n,x,k_1,t}^{k_2}) \right) \\
 & + \text{bigM} \sum_{n,c,t} (\text{slack}_{n,c,t}^+ - \text{slack}_{n,c,t}^-)
 \end{aligned} \tag{6.4}$$

**ε-constrained emissions** In light of no enforced cost of carbon in district energy systems, studies tend to look at how capping operational emissions will impact the system monetary

Table 6.2 Cost of carbon range, as given by various sources. All values given in 2017 USD/tCO<sub>2</sub>

Source	2017	2020	2030	2050
(Hourcade <i>et al.</i> , 2017)	N/A	5 - 40	N/A	105 - 130
(BEIS, 2018b)	0 - 6	0 - 13	24 - 92	N/A
(IWG, 2016)	N/A	14 - 73	19 - 86	31 - 112

cost. To do so, the constraint given in Equation 6.5 is added to a model which has cost minimisation (Equation 6.1) as its objective. The baseline emissions in Equation 6.5 are taken from the solution of the cost minimisation when an emissions cap is ignored. As with  $\epsilon$ -constrained investment cost, the scale factor  $\epsilon$  can be varied to produce a range of results, from the cost-minimised emissions to the carbon-minimised emissions. From these results, an effective cost of carbon can be calculated from the reduction in carbon emissions and the impact that has on the OFV compared to its value when carbon emissions are ignored.

$$\sum_{n,x,c,t} E_{n,x,c,t}^- \times cost_{n,x,k_2,t}^{con} \leq \epsilon \times cost_{k_2}^{op,baseline} \quad (6.5)$$

### Application of the objective function

There is only a small large range in possible carbon emissions, between the baseline carbon and baseline cost models. Table 6.3 shows that the Cambridge district could realise only a 20% reduction in emissions, and the Bangalore district only 25%. Although small, the extremes in carbon emissions are used as the minimum and maximum system emissions for  $\epsilon$ -constrained emissions analysis. This leads to an  $\epsilon$  range of 0.8–1 for Cambridge and 0.75–1 for Bangalore, where 1 refers to the carbon emissions of the baseline cost model. The pareto fronts in Figure 6.8 describe the optimal cost system for varying carbon emission limits. The shape is in agreement with previous studies: there is a steep decline in emissions for a relatively low cost increase, but the final reduction of emissions is not possible without a considerable cost increase. Centralised technologies are quickly replaced by building-level technologies in both systems as the carbon emission limit is relaxed, seen in Figure 6.9. To reach the lowest possible carbon emission levels, a combination of a large range of technologies is key. In Cambridge, every available technology in the district is used to some extent at  $\epsilon = 0.8$  and reliance on centralised energy continues until  $\epsilon = 0.88$ . Although DGs have a lower emissions factor than GridE in the Bangalore district, the marginal difference between the two leads to only having DGs in a system which has to reach minimum carbon levels. Similarly, coupling the B-CCHP with large thermal energy storage (StoreT) capacity is only necessary in the  $\epsilon = 0.75$  case. The coupling still exists until  $\epsilon = 0.95$ , but it is much weaker. There is a preference for CCHP operation to follow heat load, instead of using StoreT. Whereas investment in PV is consistent in Bangalore, there is a linear reduction in PV capacity in Cambridge, showing it to be heavily dependent on a carbon incentive for its deployment.

Table 6.3 Optimal investment costs, operation costs, and carbon emissions for baseline carbon and cost models in both Cambridge and Bangalore case studies.

	Cambridge			Bangalore		
	$cost^{inv}$	$cost^{op}$	$carbon^{op}$	$cost^{inv}$	$cost^{op}$	$carbon^{op}$
<b>Baseline cost model</b>	$2.40 \times 10^5$	$1.50 \times 10^7$	$7.14 \times 10^7$	$5.00 \times 10^7$	$2.89 \times 10^8$	$2.53 \times 10^7$
<b>Baseline carbon model</b>	$5.77 \times 10^8 \uparrow$	$1.17 \times 10^7 \downarrow$	$5.66 \times 10^7 \downarrow$	$2.76 \times 10^8 \uparrow$	$5.21 \times 10^8 \uparrow$	$1.90 \times 10^7 \downarrow$

The cost of carbon is also given for each dominant solution on the Pareto front. The possible range of imposed carbon costs, given previously in Table 6.2, is 0.365 INR/kgCO<sub>2</sub> - 9.49 INR/kgCO<sub>2</sub> (0.0036 GBP/kgCO<sub>2</sub> - 0.1 GBP/kgCO<sub>2</sub>). In both case studies, if the highest cost in the range were applied to the model, a 10% reduction in carbon emissions could be realised by a cost-focussed developer. A lower cost could still be applied in Bangalore, leading to a lower emissions saving. In Cambridge, the minimum cost of carbon required to have any effect is 0.7 GBP/kgCO<sub>2</sub>. The most realistic, near-term costs of carbon are an order of magnitude below this.

### 6.1.6 The carbon-focussed, cost-aware developer

Realistically, investment costs for a development will not be able to increase by the several orders of magnitude required to meet carbon minimisation objectives. Instead, carbon minimisation can be set as the objective, but with a limit on investment cost. The objective function in Equation 6.6 employs the same  $\epsilon$  constraint on investment cost as seen previously in Section 6.1.3.

$$\begin{aligned}
 \min \sum_{n,x,t} & \left( E_{n,x,t}^- \times cost_{n,x,k_2,t}^{con} \right) \\
 \text{s.t.} \sum_{n,x,y} & y_{n,x} \times cost_{n,x}^y \times \text{TSF} \leq \epsilon \times cost_{k_1}^{inv,baseline}
 \end{aligned} \tag{6.6}$$

The Cambridge Pareto front given in Figure 6.10a shows a linear variation in the minimum possible carbon emissions as the investment constraint changes. This is matched by a lack of drastic changes in technology capacity. Capacities for the Cambridge case are identical when comparing the cost-focussed developer and occupant objectives (Figure 6.11a vs 6.4a).

The Bangalore Pareto front shown in Figure 6.10b is not linear, but is certainly less convex than other Pareto fronts for the case study. It is also apparent that the highest operation costs are not associated with the lowest carbon emissions, and vice versa, unlike in Figure 6.3. In fact, the operation costs increase significantly at  $\epsilon = 0.31$ . Greater investment in DGs, StoreT and CCHP occurs as the cost constraint is relaxed, but the centralised technologies are only possible for  $\epsilon$  values above 1.2. The introduction of a district system occurred at a similar value of  $\epsilon$  when cost minimisation was the objective (Figure 6.4b), but with investment in a

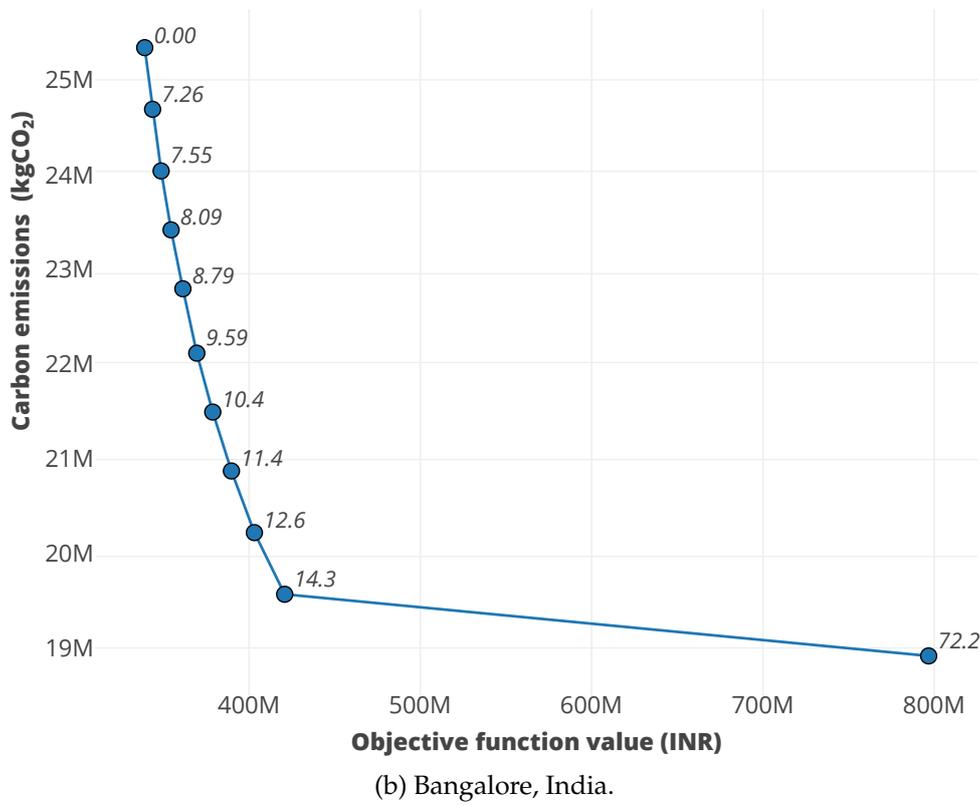
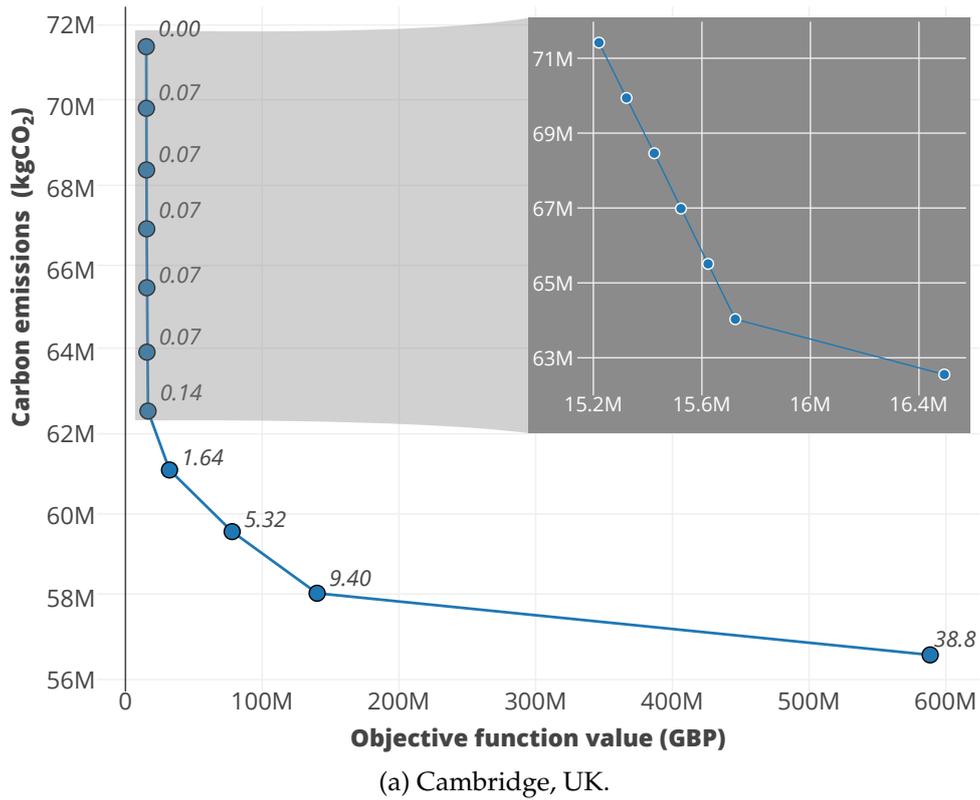


Figure 6.8 Cost-carbon Pareto front following monetary cost minimisation of both case studies with  $\epsilon$ -constrained carbon emissions.

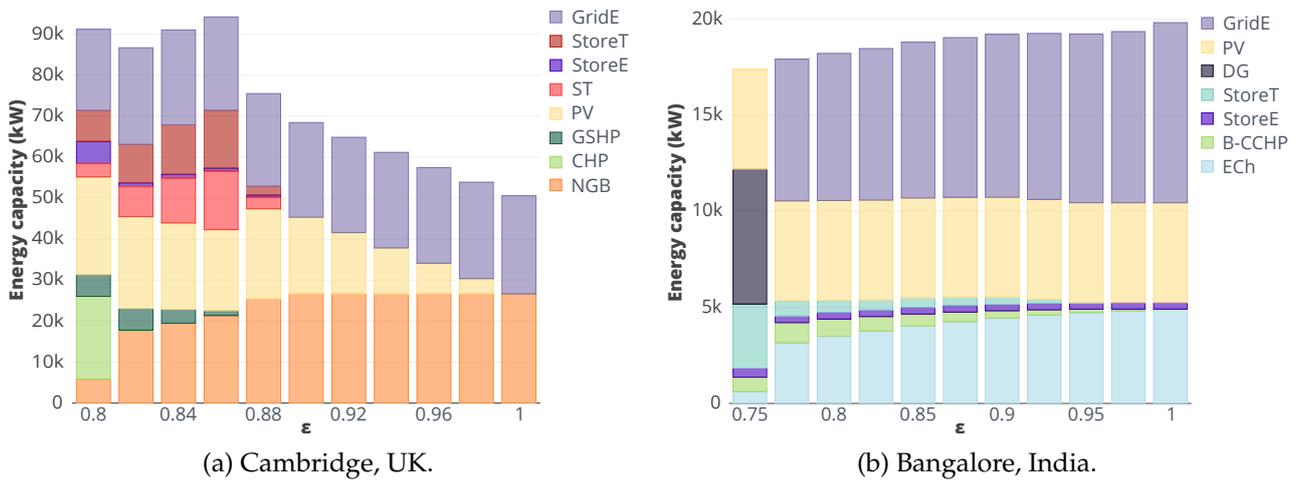


Figure 6.9 Installed energy capacity following monetary cost minimisation of both case studies with  $\epsilon$ -constrained carbon emissions.

large-scale ECh taking preference over the CCHP. In both case studies, PV capacity increases as the investment cost constraint is relaxed.

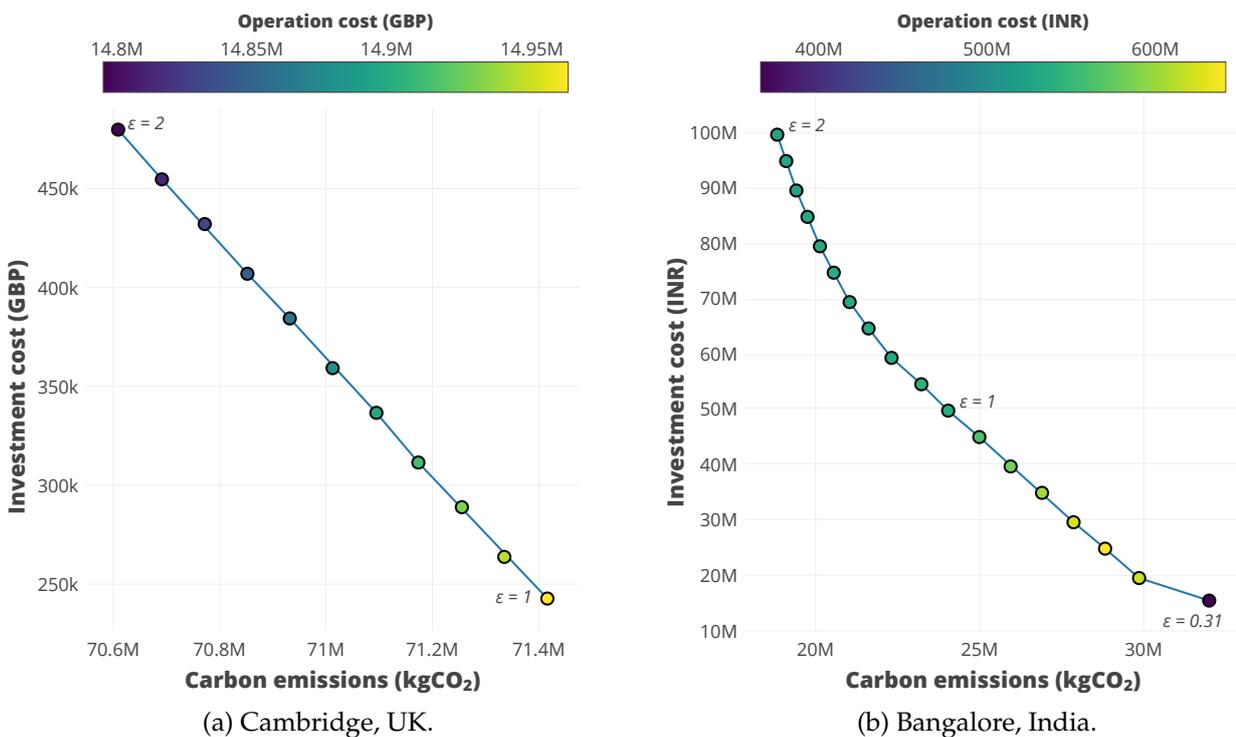


Figure 6.10 Investment cost-carbon Pareto front following carbon emission minimisation of both case studies with  $\epsilon$ -constrained investment cost. Operation cost associated with each model run is given by the marker colour.

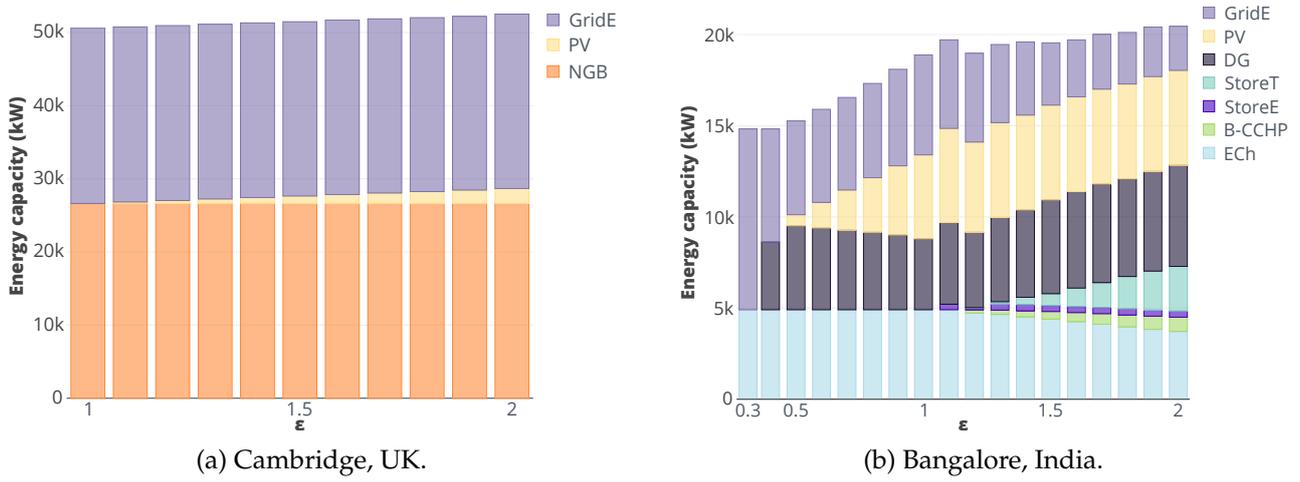


Figure 6.11 Installed energy capacity following carbon emission minimisation of both case studies with  $\epsilon$ -constrained investment cost.

### 6.1.7 Cost-focussed, district-constrained client

If there are no constraints on the provision of thermal demand, Chapter 5 concluded that building-level solutions are optimal, in both a heating and a cooling district network. However, thermal demand can be completely met by a district system. The monetary impact of this can be seen if using the constraint given in Equation 6.7. By decreasing the allowed contribution of building-level thermal energy technologies, district systems must balance system energy. It is likely that there is a step-change in cost when such a constraint is applied, since purchasing energy centre technologies requires significant up-front costs. However, it remains to be seen how costs change as the dependence on building-level solutions decreases. For instance, if building-level technologies can only meet up to 80% of thermal demand, a greater than 20% contribution from the centralised system may prove cost-optimal, given that the energy centre has already been purchased.

$$\sum_{n,x'',c,t} \mathbf{E}_{n,x'',c,t}^+ \leq \epsilon \times \sum_{n,x',c,t} \mathbf{E}_{n,x',c,t}^+ \quad (6.7)$$

Where  $x'$  = thermal energy producing technology,  $x''$  = building-level thermal energy producing technology.

Each case study model was optimised with an increasingly strict requirement to not meet thermal demand by building-level generation methods (e.g. NGBs or EChs). Optimal investment decisions are given in Figure 6.12. In Cambridge, this leads to the introduction of a CHP, coupled with a large thermal storage vessel. Due to the CHP electricity generation, the need for GridE decreases. Building-level NGBs are maintained at a relatively large capacity, to meet peak demand. As such, total NGB capacity does not reduce at the same rate as the limit on the systemwide limit on their contribution to thermal demand. For instance, at a 50% limit, the NGB capacity is not 50% of that seen where there is no limit (100%). A fully centralised heat network also requires a small amount of NGB capacity at the energy centre, to meet this same peak-load function. However, in Bangalore the electricity generating

technologies remain undisturbed and the building-level EChs are replaced by the more efficient EC-ECh. As the share of building-level provision drops below 30%, StoreT at the energy centre becomes financially viable. As such, the EC-ECh has a higher capacity factor at low building-level provision, with StoreT storing energy in time periods of low thermal demand. A small CCHP is also added to the investment portfolio as soon as a centralised system is required (90% building level share), but its capacity does not increase noticeably as the energy centre becomes more prominent.

As the share of building-level provision decreases, not all buildings in the network are provided for equally. Table 6.5 shows the change in thermal demand met by building-level technologies at each node, as the limit is reduced on the sum of district thermal demand met by building-level technologies. It can be seen, by the inconsistency in colour change between rows, that some buildings in the network remain independent from the centralised system longer than others. For instance, building *B* in the Bangalore district has 100% of its demand met by a building-level ECh until the whole district is forced to meet less than 70% of its demand from building-level technologies. Buildings *J* and *F* take the slack, with consistently lower demand met by building-level EChs (table values) than the district requirement (column values). The difference between buildings in the network is more apparent in Cambridge; many buildings depend only on centralised heat (table values = 0) well before the full system is required to do so (column value = 0). But certain buildings, namely *B02* and *B07*, depend relatively heavily on building-level NGBs up until the whole system has to be centralised. Cross-over points are also apparent, such as *A17*, which depends on its building-level NGB for more than 90% of its heat demand until the system is required to meet more than 50% of its demand using centralised heat provision. At that point, the building rapidly decreases its dependence on its own NGB.

The relative difference between nodes on the network is an expected result of their distance from the energy centre. Indeed, the greater variation between nodes in the Cambridge district is due to its larger size relative to the Bangalore district. The nodes in the *D* and *C* zones sign up to the district network first. Nodes *B01*, *B02*, *B03*, *B07*, *B08* are closer to the energy centre than *A01*, *A02*, *A05a*, *A08*, if taking a straight line distance. However, road following for the heat network leads to the two sets of nodes being equally distant from the energy centre ( $\sim 1.2\text{km}$ ). The furthest building from the energy centre is *A08* ( $\sim 1.25\text{km}$ ), but more than 50% of its heat demand is met by the centralised heat network well before the constraint on system heat demand met by building-level technologies reaches 50%. This is primarily due to its relatively low heat demand, e.g. compared to the preceding building in the network, *A07*.

The investment cost incurred by the centralised system is significant in the Cambridge case; an order of magnitude increase in investment cost is required to meet 10% of the district demand using a centralised system (Table 6.4a). This cost increases by a further two orders of magnitude up to a fully centralised thermal system. Although there is also a high cost for a fully centralised system in the Bangalore, the cost is almost proportional to the increased prevalence of the centralised system (Table 6.4a). For instance, it costs 40% more to invest in a system which meets 40% of district demand with a centralised system. The change in carbon emissions and operation cost is not noticeable, relative to the change in investment

cost. Nevertheless, carbon emissions decrease with a centralised system in Bangalore, but increase with one in Cambridge.

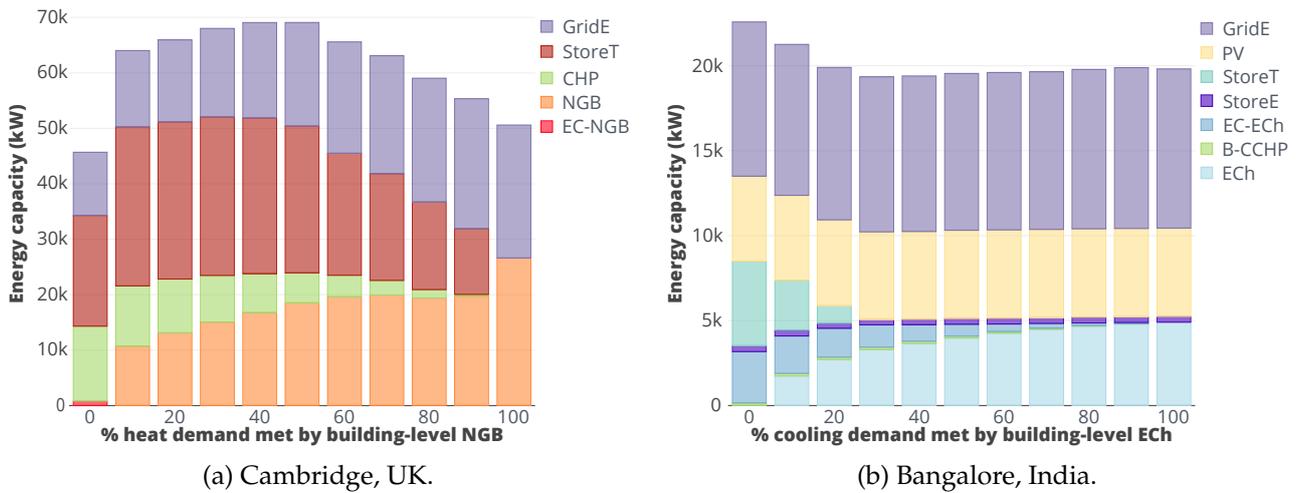


Figure 6.12 Installed energy capacity following monetary cost minimisation of both case studies with a percentage limit set on the sum of district thermal demand which can be met by building-level conventional technologies.

Table 6.4 Monetary cost and carbon emissions following monetary cost minimisation of both case studies with a percentage limit set on the sum of district thermal demand which can be met by building-level conventional technologies.

	Share of heat demand met by building-level NGB										
	0	10	20	30	40	50	60	70	80	90	100
Operation (GBP)	1.1x10 <sup>7</sup>	1.1x10 <sup>7</sup>	1.2x10 <sup>7</sup>	1.2x10 <sup>7</sup>	1.3x10 <sup>7</sup>	1.3x10 <sup>7</sup>	1.4x10 <sup>7</sup>	1.4x10 <sup>7</sup>	1.4x10 <sup>7</sup>	1.5x10 <sup>7</sup>	1.5x10 <sup>7</sup>
Carbon (kgCO <sub>2</sub> )	7.4x10 <sup>7</sup>	7.2x10 <sup>7</sup>	7.3x10 <sup>7</sup>	7.4x10 <sup>7</sup>	7.5x10 <sup>7</sup>	7.5x10 <sup>7</sup>	7.4x10 <sup>7</sup>	7.4x10 <sup>7</sup>	7.4x10 <sup>7</sup>	7.3x10 <sup>7</sup>	7.1x10 <sup>7</sup>
Investment (GBP)	3.5x10 <sup>8</sup>	2.1x10 <sup>8</sup>	1.5x10 <sup>8</sup>	1.1x10 <sup>8</sup>	7.8x10 <sup>7</sup>	4.9x10 <sup>7</sup>	3.0x10 <sup>7</sup>	1.6x10 <sup>7</sup>	8.1x10 <sup>6</sup>	2.0x10 <sup>6</sup>	2.4x10 <sup>5</sup>

(a) Cambridge, UK.

	Share of cooling demand met by building-level ECh										
	0	10	20	30	40	50	60	70	80	90	100
Operation (INR)	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>	2.8x10 <sup>8</sup>
Carbon (kgCO <sub>2</sub> )	2.4x10 <sup>7</sup>	2.4x10 <sup>7</sup>	2.4x10 <sup>7</sup>	2.4x10 <sup>7</sup>	2.4x10 <sup>7</sup>	2.4x10 <sup>7</sup>	2.4x10 <sup>7</sup>	2.4x10 <sup>7</sup>	2.4x10 <sup>7</sup>	2.5x10 <sup>7</sup>	2.5x10 <sup>7</sup>
Investment (INR)	2.3x10 <sup>8</sup>	1.5x10 <sup>8</sup>	1.2x10 <sup>8</sup>	1.0x10 <sup>8</sup>	9.1x10 <sup>7</sup>	7.9x10 <sup>7</sup>	7.0x10 <sup>7</sup>	6.3x10 <sup>7</sup>	5.7x10 <sup>7</sup>	5.3x10 <sup>7</sup>	5.0x10 <sup>7</sup>

(b) Bangalore, India.

Table 6.5 Percentage share of thermal demand met by building-level thermal technologies in each node of both case study districts. Each column gives results for a different limit on the district-wide thermal demand that can be met by building-level technologies. Each row gives the percentage share of thermal demand met by building-level technologies in a particular node. Lower values (dark green) highlight the nodes which depend more heavily on centrally-produced thermal energy. Conversely, higher values (light yellow) depend more heavily on building-level thermal energy. For the location of each node, see Figure 5.5 on page 100.

		% share of heat demand met by building-level NGB										
		0	10	20	30	40	50	60	70	80	90	100
Nodes	A01	0	36.8	66.6	66.6	66.6	66.7	66.7	66.7	66.9	67.1	100
	A02	0	19.1	67.8	68	68.1	68.1	68.1	68.2	68.2	68.5	100
	A03	0	19	37.7	49.7	49.7	49.7	49.8	49.9	50	55.6	100
	A04	0	0	0	18.9	47.2	47.3	47.4	47.4	47.5	56.8	100
	A05a	0	31.5	64.3	83.9	87.4	87.4	87.4	87.4	87.5	87.6	100
	A05b	0	0	4.8	44.1	55.5	55.5	55.5	55.5	55.6	58.1	100
	A07	0	20.1	41.2	71.3	94	94	94	94	94.1	94.1	100
	A08	0	33.7	47.2	47.3	47.3	47.3	47.5	47.5	47.5	56	100
	A09a	0	0	0	13.3	46.2	46.3	46.3	46.3	46.5	57.2	100
	A09b	0	0	28.1	46	46	46	46	46.1	46.2	58.3	100
	A10	0	0	0	5.2	50.4	50.5	50.5	50.5	50.7	56.8	100
	A10b	0	0	0	0	36.2	67.7	67.7	67.8	67.9	68.2	100
	A11	0	2.8	9.6	51.2	51.2	51.2	51.3	51.3	51.4	58.1	100
	A12	0	15.3	38.8	56.9	56.9	56.9	56.9	56.9	57	58.6	100
	A13	0	18.9	21.8	46.6	46.6	46.6	46.6	46.6	47	55.8	100
	A14	0	0	0	0	0	56.2	56.2	56.2	56.2	57.5	100
	A17	0	3.2	10.5	29.1	59.6	92.9	92.9	92.9	93	93	100
	A18	0	4.6	13.4	29.6	49.2	49.2	49.3	49.4	49.5	55.4	100
	A19	0	0	0	0	18.5	47.7	47.9	47.9	47.9	57.4	100
	A20	0	0	0	0	0	47.4	47.6	47.6	47.6	56.9	100
A21	0	1.5	1.8	2.9	3.8	22.3	78	81.5	81.6	81.7	100	
A22	0	0	0	0	0	0	35.9	56	56.1	58.4	100	
B01	0	26.7	48	71.2	83.5	86.9	86.9	87	87	87.1	100	
B02	0	40.3	49.7	70.3	70.3	70.3	70.3	70.4	70.5	70.7	100	
B04	0	0	0	0	0.1	65.2	65.2	65.2	65.4	65.8	100	
B05a	0	0	0	0	0	0.4	52.7	52.8	52.9	56.7	100	
B06	0	0	0	0	4.4	11.1	42.2	75.7	75.7	76.4	100	
B07	0	42.5	69.1	70.3	70.4	70.4	70.4	70.5	70.6	71	100	
B08	0	28.6	46.2	46.3	46.3	46.3	46.3	46.4	46.7	55.3	100	
C01	0	0	0	0	0	0	0	5	71.7	72.3	100	
C02	0	0	0	0	0	0	0	0	0	55.8	100	
C04	0	0	0	0	0	0	4.1	20.1	55.9	83.6	100	
C07a	0	0	0	0	0	0	1.7	55.1	55.2	56.1	100	
C07b	0	6.6	6.6	4.1	4.1	7.7	24.9	73.9	73.9	74.6	100	
C08	0	0	0	0	0	0	0	0	47.2	55.8	100	
D01	0	0	0	0	0	47.6	47.7	47.7	47.7	57.6	100	
D02	0	0	0	0	0	0	0	40.4	64.4	64.4	100	
D04	0	0	0	0	0	0	0	0	0	23.9	100	
D06	0	0	0	0	0	0	0	0	0	23.5	100	

(a) Cambridge, UK.

		% share of cooling demand met by building-level ECh										
		0	10	20	30	40	50	60	70	80	90	100
Nodes	A	0	21.5	27.3	50.1	62.9	75.4	79.3	84.9	100	100	100
	B	0	23.2	36.2	57.7	71.8	78.1	82.7	100	100	100	100
	C	0	19.6	24.8	38.8	55.7	68.8	75.3	80.6	100	100	100
	D	0	21.4	27.2	50.2	62.7	75.3	78.9	84.8	100	100	100
	E	0	4.7	20	24.5	31.7	49.5	57.8	69.4	77.1	89.7	100
	F	0	3.5	16.9	23.8	26.5	38.6	50.7	59.9	73.5	81.3	100
	G	0	14.8	23.5	28.9	48.5	59.6	70.1	76.9	83.1	100	100
	H	0	13.8	23.4	28.9	47.6	58.7	69.9	76.8	82.9	100	100
	I	0	22	27.6	52	64	75.7	79.9	89.2	100	100	100
	J	0	2.4	12.2	22.2	25	31.5	43.1	55.9	69	78.1	100
	K	0	13.6	23.1	27.8	44.7	58.6	68.8	76.3	81.8	100	100

(b) Bangalore, India.

### 6.1.8 Risk-aware client

Reformulations of the objective function have hitherto remained deterministic. It was shown in Chapter 5 that deterministic models suffer from instability when subjected to unexpected future demand profiles and power interruptions. To improve system robustness, SO methodology is extended in this chapter to reduce the risk of future scenarios incurring high financial penalties and to prevent the worst-case carbon emissions from being realised. A cost of carbon baseline model using only the mean demand scenario is used for comparison. Risk-awareness has only been introduced to the Bangalore case study, to explore the possibility of its introduction when carbon emissions are accounted for in a district.

#### Objective function formulation

The risk-aware objective function when minimising monetary cost is given again in Equation 6.8. Each scenario is probabilistically weighted, with the highest risk scenarios further penalised using the **conditional value at risk (CVaR)** component. When mitigating risk for the carbon-focussed client, only the **VaR** is scenario-independent in the objective function (Equation 6.9). Investment decisions must still be made to satisfy all scenarios, but, in agreement with the deterministic objective function, the embodied carbon of the investment is not a consideration. Scenarios which may incur high carbon emissions can be avoided by imposing a monetary cost on carbon emissions (i.e. the cost of carbon). By including only the cost of carbon within the **CVaR** component of Equation 6.10, system cost can be minimised across multiple scenarios while ensuring high emission risk remains manageable. It would be feasible in this objective function to use the carbon emissions directly in the **CVaR** component, but using the cost of carbon allows us to appropriately scale the emissions to match the magnitude of monetary costs. A 9.49 INR/kgCO<sub>2</sub> cost of carbon has been applied, representing the higher end of possible costs given in Table 6.2.

$$\begin{aligned}
\min \quad & TSF \times \sum_{n,x,y} \left( \mathbf{y}_{n,x} \times \text{cost}_{n,x,k}^y \right) \\
& + \sum_{n,x,c,s',t} W_{s'} \times \left( \mathbf{E}_{n,x,c,s',t}^+ \times \text{cost}_{n,x,k_1,t}^{prod} - \mathbf{E}_{n,x,c,s',t}^- \times \text{cost}_{n,x,k_1,t}^{con} \right. \\
& \quad \left. + \mathbf{E}_{n,x,c,s',t}^{ex} \times \text{cost}_{n,x,k_1,t}^{ex} \right) + \text{bigM} \sum_{n,c,s',t} W_{s'} \times \left( \mathbf{slack}_{n,c,s',t}^+ - \mathbf{slack}_{n,c,s',t}^- \right) \\
& + \beta \left( \zeta + \frac{1}{1-\alpha} \sum_{s'} (W_{s'} \eta_{s'}) \right) \\
\text{s.t.} \quad & \sum_{n,x,c,s',t} W_{s'} \left( \mathbf{E}_{n,x,t}^+ \times \text{cost}_{n,x,k_1,t}^{prod} - \mathbf{E}_{n,x,c,s',t}^- \times \text{cost}_{n,x,k_1,t}^{con} + \mathbf{E}_{n,x,c,s',t}^{ex} \times \text{cost}_{n,x,k,t}^{ex} \right) \\
& + \text{bigM} \sum_{n,c,s',t} W_{s'} \times \left( \mathbf{slack}_{n,c,s',t}^+ - \mathbf{slack}_{n,c,s',t}^- \right) - \zeta \leq \eta_{s'}
\end{aligned} \tag{6.8}$$

$$\begin{aligned}
\min \quad & \sum_{n,x,c,s',t} W_{s'} \times \left( -\mathbf{E}_{n,x,c,s',t}^- \times \text{cost}_{n,x,k_2,t}^{\text{con}} \right) \\
& + \text{bigM} \sum_{n,c,s',t} W_{s'} \times \left( \text{slack}_{n,c,s',t}^+ - \text{slack}_{n,c,s',t}^- \right) \\
& + \beta \left( \xi + \frac{1}{1-\alpha} \sum_{s'} (W_{s'} \eta_{s'}) \right) \\
\text{s.t.} \quad & \sum_{n,x,c,s',t} W_{s'} \left( -\mathbf{E}_{n,x,c,s',t}^- \times \text{cost}_{n,x,k_2,t}^{\text{con}} \right) \\
& + \text{bigM} \sum_{n,c,s',t} W_{s'} \times \left( \text{slack}_{n,c,s',t}^+ - \text{slack}_{n,c,s',t}^- \right) - \xi \leq \eta_{s'}
\end{aligned} \tag{6.9}$$

$$\begin{aligned}
\min \quad & \text{TSF} \times \sum_{n,x,y} \left( \mathbf{y}_{n,x} \times \text{cost}_{n,x,k}^y \right) \\
& + \sum_{n,x,c,s',t} W_{s'} \times \left( \mathbf{E}_{n,x,c,s',t}^+ \times \text{cost}_{n,x,k,t}^{\text{prod}} + \mathbf{E}_{n,x,c,s',t}^{\text{ex}} \times \text{cost}_{n,x,k,t}^{\text{ex}} \right. \\
& \left. - \mathbf{E}_{n,x,c,s',t}^- \times \left( \text{cost}_{n,x,k,t}^{\text{con}} + \text{cost}_{n,x,k_2,t}^{\text{con}} \times \text{cost}_{n,x,k_1,t}^{k_2} \right) \right) \\
& + \text{bigM} \sum_{n,c,s',t} W_{s'} \times \left( \text{slack}_{n,c,s',t}^+ - \text{slack}_{n,c,s',t}^- \right) \\
& + \beta \left( \xi + \frac{1}{1-\alpha} \sum_{s'} (W_{s'} \eta_{s'}) \right) \\
\text{s.t.} \quad & \sum_{n,x,c,s',t} W_{s'} \left( -\mathbf{E}_{n,x,c,s',t}^- \times \text{cost}_{n,x,k_2,t}^{\text{con}} \times \text{cost}_{n,x,k_1,t}^{k_2} \right) \\
& + \text{bigM} \sum_{n,c,s',t} W_{s'} \times \left( \text{slack}_{n,c,s',t}^+ - \text{slack}_{n,c,s',t}^- \right) - \xi \leq \eta_{s'}
\end{aligned} \tag{6.10}$$

### Scenario reduction

As discussed in Section 5.1.2, reduced scenarios are selected using the objective function values from optimising each of the 500 scenarios independently. Figure 6.13 shows the comparison of applying a minimum cost and a minimum carbon objective function to each of the scenarios. There is a clear correlation between the two: a scenario which incurs a higher optimal cost will incur higher optimal carbon emissions. In selecting reduced scenarios, the scenarios in Figure 6.13 are subset horizontally in preparation for cost minimisation SO and vertically in preparation for carbon minimisation SO. For cost minimisation SO including a cost of carbon, scenario reduction (SR) is applied to the sum of normalised cost and carbon minimisation OFVs of each scenario.

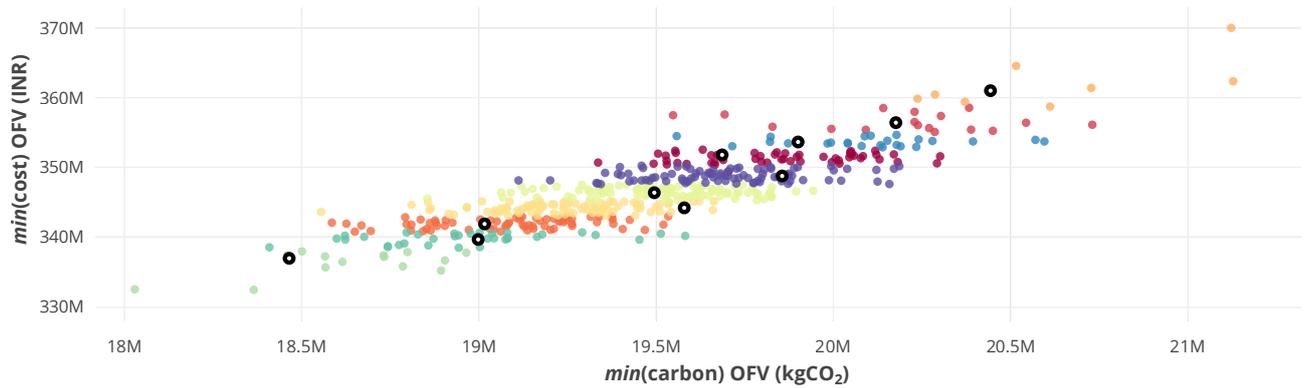
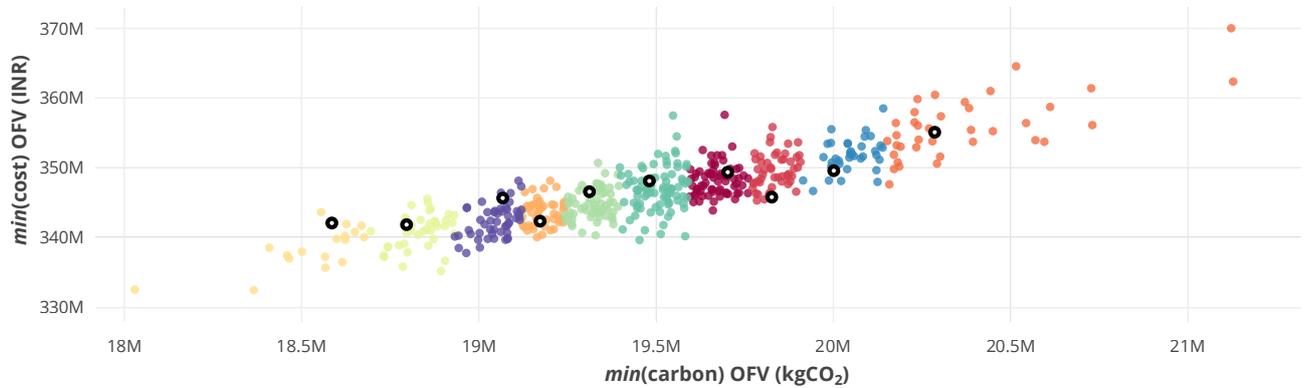
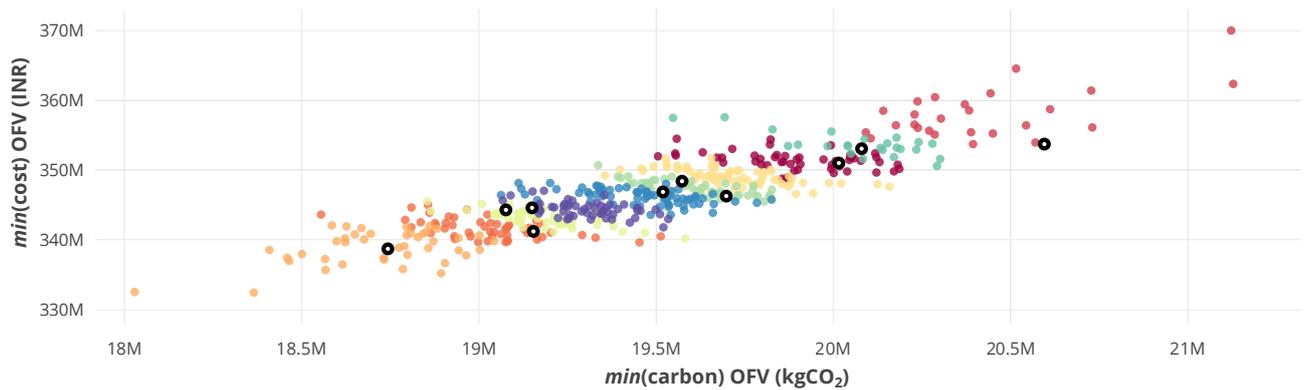
(a) SR using  $\min(\text{cost})$  OFV.(b) SR using  $\min(\text{carbon})$  OFV.(c) SR using normalised sum of  $\min(\text{cost})$  and  $\min(\text{carbon})$  OFVs.

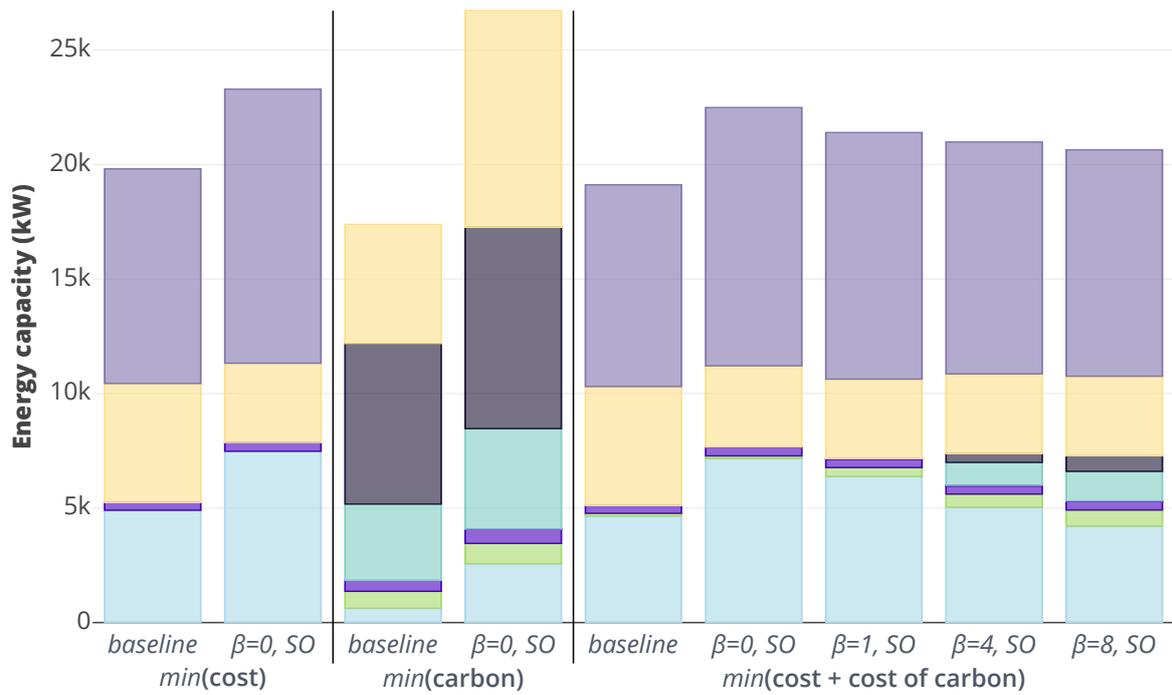
Figure 6.13 Cost-optimal vs. carbon-optimal OFV for 500 generated scenarios in the Bangalore case study. The same scatter is shown overlaid with application of SR to (a) cost minimisation OFVs, (b) carbon emissions minimisation OFVs and (c) sum of normalised cost and carbon minimisation OFVs. Reduced scenarios, chosen to represent the full scenario set, are highlighted with a thick black outline. All scenarios represented by a particular reduced scenario have the same marker colour.

### Scenario optimisation

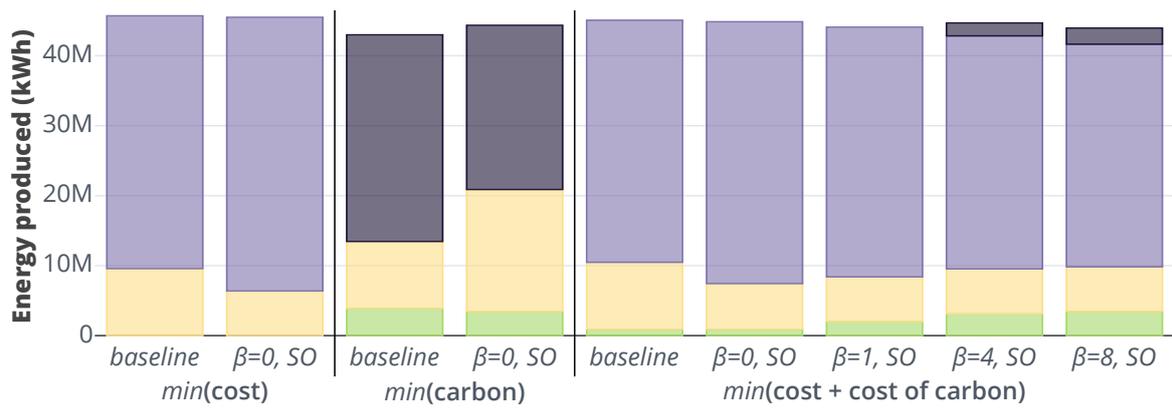
The optimal investment portfolios and average annual energy production following **SO** are given in Figure 6.14. As found in Chapter 5, the introduction of several scenarios in the objective function leads to an increase in technology capacity and a decrease in expected capacity factor (annual energy production or technology capacity). Increasing capacity is straightforward in the carbon minimisation **SO** model since there is no carbon cost associated with the additional investment. Hence, **PV** and **ECh** capacity doubles. Conversely, when cost is part of the objective function, **ECh** capacity increases by 50% but **PV** capacity reduces.

Introducing carbon emissions into the cost minimisation objective, using a cost of carbon, leads to consistent investment in a centralised **CCHP**. This occurs in the baseline, single scenario model and the **SO** models. The capacity is low when neutral to risk, increasing as the risk increases. Although the technology capacities would suggest that the **CCHP** only contributes a small amount to meeting demand, the low electrical efficiency and subsequent high cooling **coefficient of performance (COP)** of 2.1 leads to it contributing to most of the cooling demand at high levels of risk aversion. Cooling demand is met similarly when optimising the carbon minimisation **SO** model, but there is a stark difference in the choice of electricity generator, with cost concerns leading to heavy dependence on **GridE** instead of **PVs** and **DGs**.

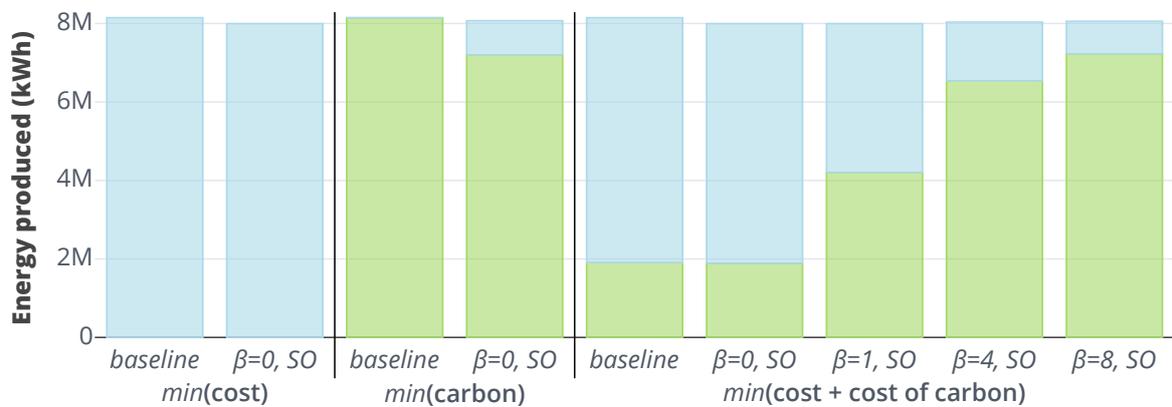
Accounting for the risk of high carbon emissions in cost minimisation **SO** leads to inevitably increased monetary costs. Figure 6.15a shows that not only does the cost increase with risk aversion, but the standard deviation of realisable costs also increases. Conversely, increased risk aversion leads to lower emitted carbon and a lower standard deviation on the distribution of realisable carbon emissions (Figure 6.15b). It is possible to reduce carbon emissions to correspond with the tail carbon emissions realised in the 500 independently optimised scenarios without significantly increasing system monetary cost. This would require a risk aversion of  $\beta = 1$ . Arguably, a risk aversion of  $\beta = 4$  offers a good compromise as it is closer to the  $\beta = 8$  carbon savings whilst sitting centrally between the  $\beta = 1$  and  $\beta = 8$  cost increase. However, through risk aversion alone it is not possible to reach the system carbon emissions achieved when setting an  $\epsilon$  constraint on investment costs. The 9.49 **INR/kgCO<sub>2</sub>** cost of carbon used here realised system carbon emissions closer to 22 million kgCO<sub>2</sub> when considering the cost of carbon in a single-scenario model (Figure 6.8b).



(a) Energy capacity.



(b) Annual electricity production.



(c) Annual cooling production.

Legend: StoreE, StoreT, PV, B-CCHP, ECh, DG, GridE

Figure 6.14 Comparison of optimal technology capacity and annual energy production for Bangalore case study baseline and SO models.

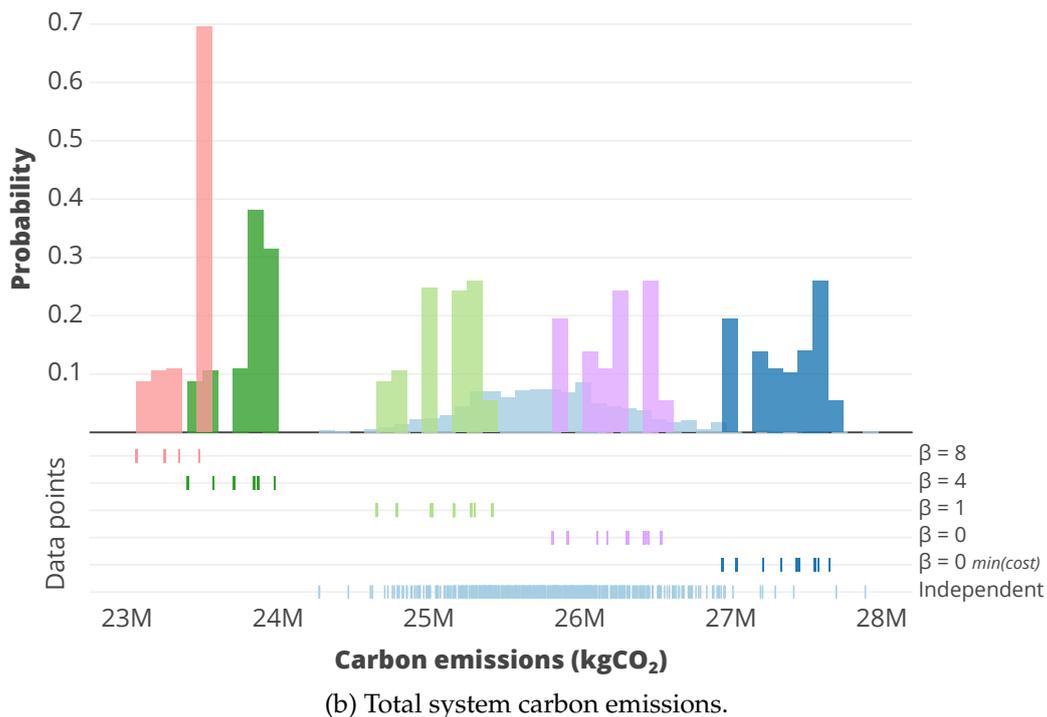
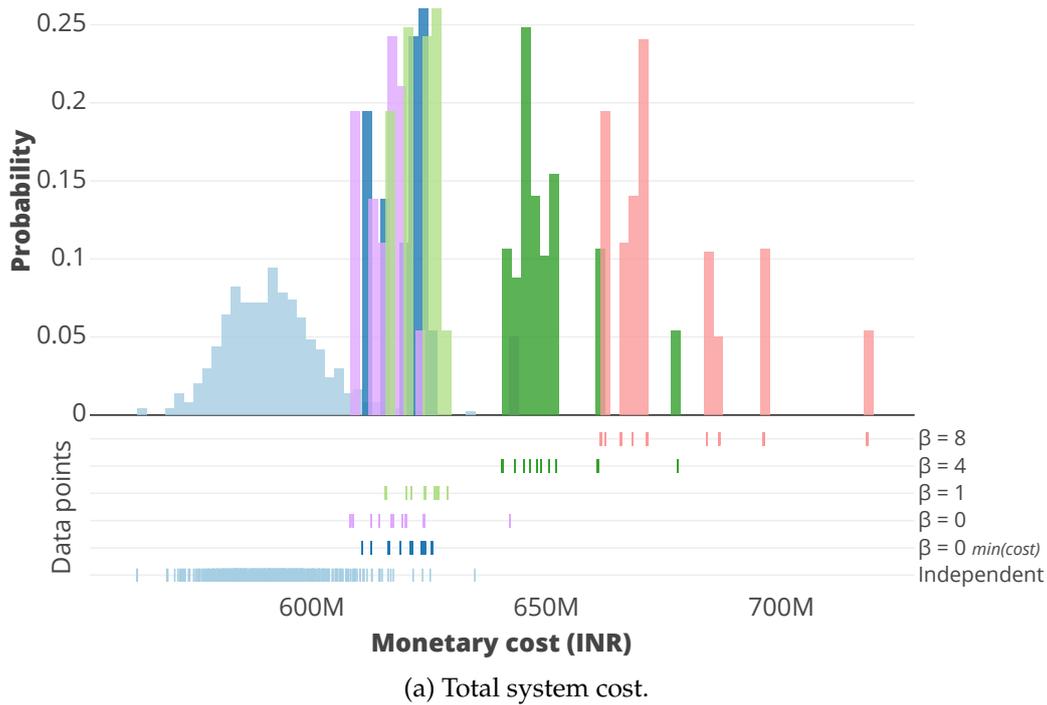


Figure 6.15 Distribution of total system monetary cost and carbon emissions resulting from minimising cost (including the cost of carbon), using risk-neutral SO and risk-averse SO Bangalore case study models. Comparison is made to risk-unaware optimisation on all 500 independent scenarios and risk-neutral SO with only cost minimisation in the objective. Data points from which the distributions are constructed are identified.

## 6.2 Out-of-sample tests

The investment portfolios resulting from different decision-maker objectives applied to the Bangalore case study district have been tested in this chapter against the same OOS tests as used in Chapter 5. Each test is a rolling horizon optimisation over a demand profile from 500 generated demand scenarios (see Section 5.4 for more information). This section begins by assessing the inadvertent impact of different single scenario objectives on operational unmet demand. The impact of incorporating demand scenarios into the objective function using both risk-neutral and risk-averse SO is then assessed, to understand whether a cost-carbon-risk objective is viable.

### 6.2.1 Single scenario models

Between single scenario models, there tends to be little variation in unmet demand, certainly when compared to the orders of magnitude variation seen between some models in the previous chapter. However, variation does exist. For instance, the imposition of a lower investment cost leads to a lower level of unmet electricity demand, when minimising monetary cost (Upper trace in Figure 6.16). The optimal investment portfolio following minimisation of carbon emissions with  $\epsilon$ -constrained investment costs (Figure 6.17) proves more robust to unmet demand than the same  $\epsilon$  constraint applied to a cost minimisation model (Figure 6.16). This may be explained by the increased StoreT capacity in the former technology portfolio, alongside the mix of DG and GridE capacity to meet electricity demand. Trends such as this are hardly strict, however, given that the baseline cost model ( $\epsilon = 1$ ) deviates considerably in unmet demand from its nearest neighbours, usually resulting in greater unmet demand. Indeed, in Figures 6.16 and 6.19, the  $\epsilon = 1$  model incurs a greater unmet cooling demand when compared to other  $\epsilon$  values. In Figure 6.17, the same poor performance is seen for unmet electricity demand.

Generally, there is lower unmet cooling demand in a district with more dependence on centralised cooling (Lower trace in Figure 6.18) and lower carbon emissions (Lower trace in Figure 6.19). This benefit is realised by the existence of StoreT capacity, which is not chosen in the baseline cost model. However, there is an increase in unmet electricity demand as carbon emissions reduce (Upper trace in Figure 6.19), indicating the lack of flexibility in the DGs chosen to meet most of the electricity demand in the carbon-focussed client model. Ultimately, there is a perceptible variation in unmet demand depending on the objective of the decision maker, but it cannot be predicted prior to running OOS tests.

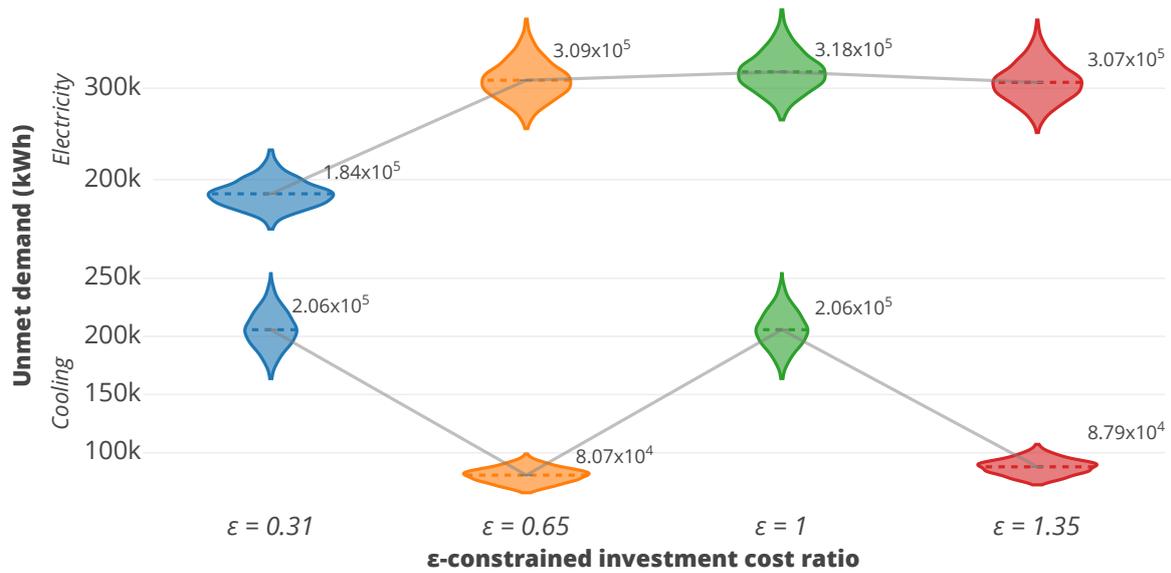


Figure 6.16 Distribution of unmet cooling and electricity demand realised when running OOS tests on the result of cost minimisation with  $\epsilon$ -constrained investment cost in the Bangalore case study district. Values adjacent to distributions refer to their mean.

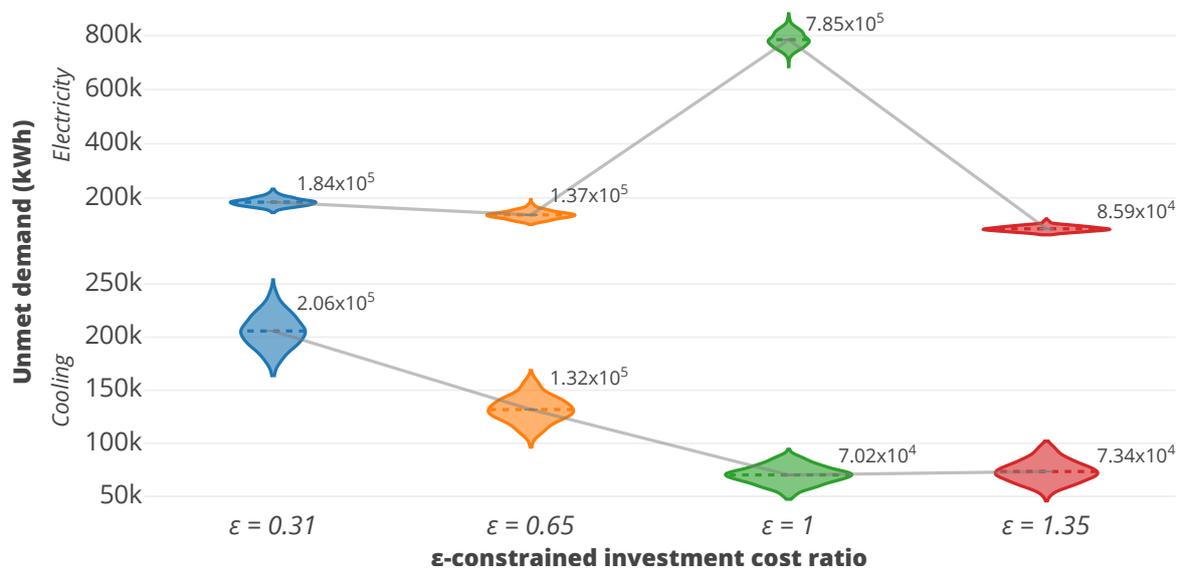


Figure 6.17 Distribution of unmet cooling and electricity demand realised when running OOS tests on the result of carbon minimisation with  $\epsilon$ -constrained investment cost in the Bangalore case study district. Values adjacent to distributions refer to their mean.

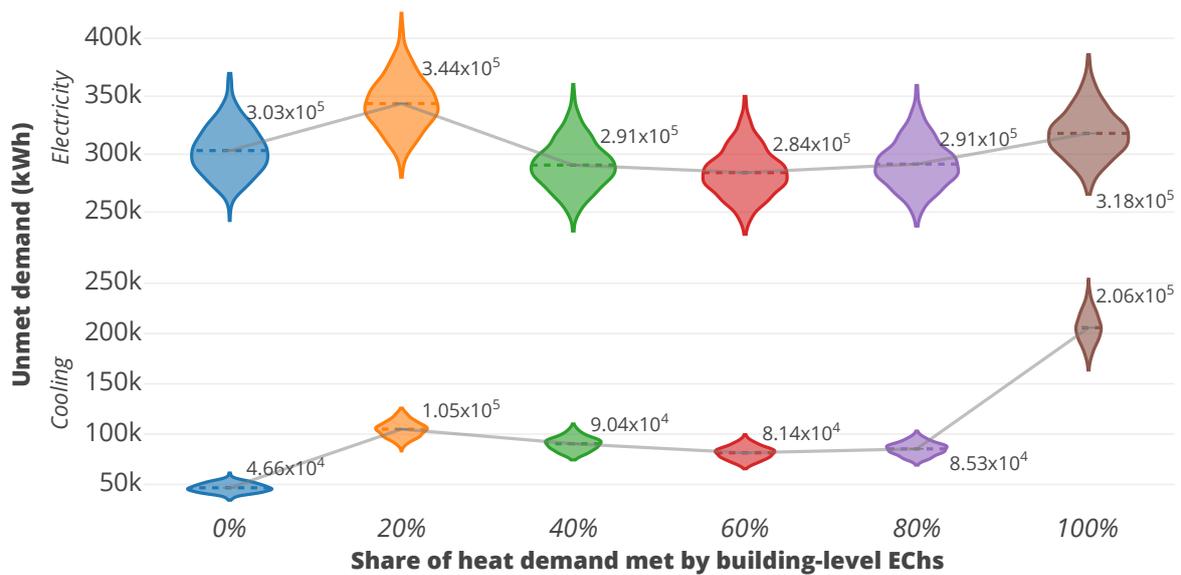


Figure 6.18 Distribution of unmet cooling and electricity demand realised when running OOS tests on the result of cost minimisation with decreasing allowed share of cooling demand met by EChs in the Bangalore case study district. Values adjacent to distributions refer to their mean.

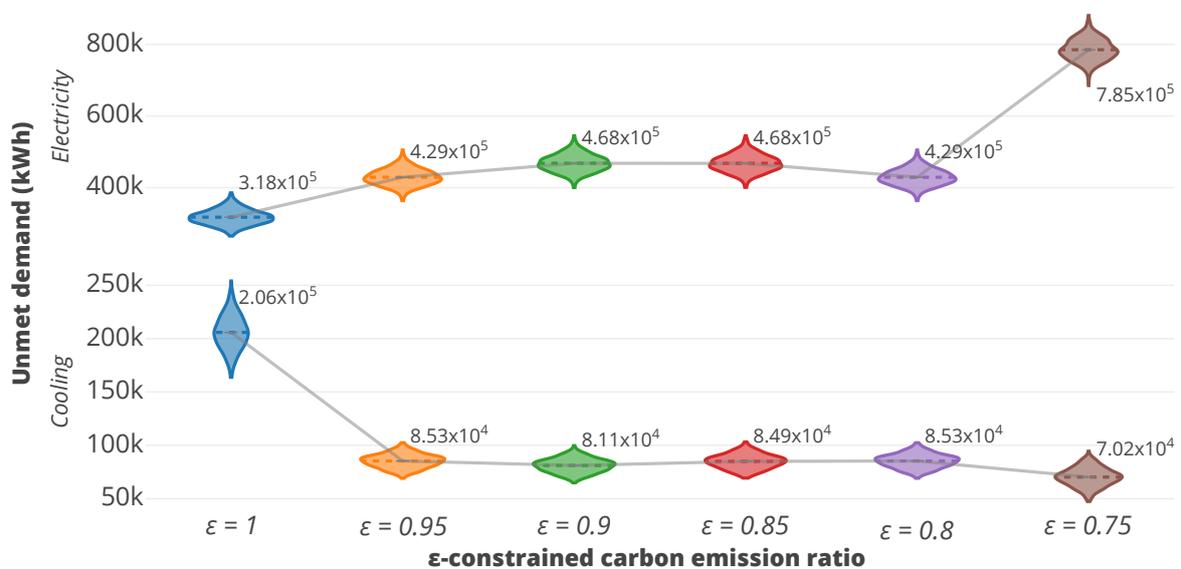


Figure 6.19 Distribution of unmet cooling and electricity demand realised when running OOS tests on the result of cost minimisation with ε-constrained carbon emissions in the Bangalore case study district. Values adjacent to distributions refer to their mean.

## 6.2.2 Scenario optimisation models

As expected, based on the results from Chapter 5, the introduction of demand scenarios inherent in the objective function of a model does drastically improve robustness to unmet demand. The baseline models are all likely to incur a one or two magnitude greater quantity of unmet demand than the baseline models. There is a similar difference between risk-neutral SO models, with the carbon minimisation objective leading to a lower unmet cooling demand than the cost minimisation objective, and vice versa for unmet electricity demand. But, the carbon minimisation objective SO model has a larger standard deviation from the 500 OOS tests than any other model seen in this chapter, or the previous chapter. Indeed, in the worst case OOS test, the SO optimal investment would lead to greater unmet electricity demand than some of the baseline models. The probability is particularly low, but it is still a possibility that is not exhibited in the other SO model results.

Applying a cost of carbon to the cost minimisation objective leads to an unmet demand range that sits between the cost and carbon minimisation results. The range of unmet demand remains low, unlike with the carbon minimisation model. Updating the CVaR risk measure to be the cost and carbon does aid in reducing the risk, compared to the risk-neutral case (Figure 6.21). Although this does not lead to the lowest unmet cooling or electricity demand, it is possible to avoid a large swing in unmet demand between the two energy types by incorporating the cost of carbon.

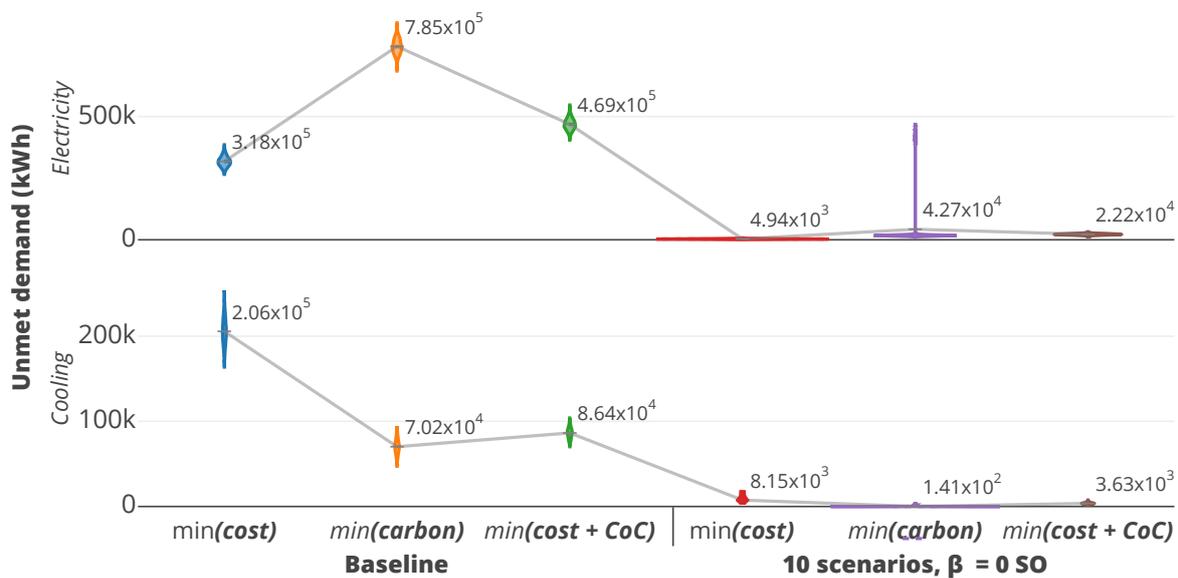


Figure 6.20 Distribution of unmet cooling and electricity demand realised when running OOS tests on the result of baseline, single scenario optimisation and risk-neutral, 10-scenario SO in the Bangalore case study district. CoC = cost of carbon,  $\beta$  = risk aversion measure. Values adjacent to distributions refer to their mean.

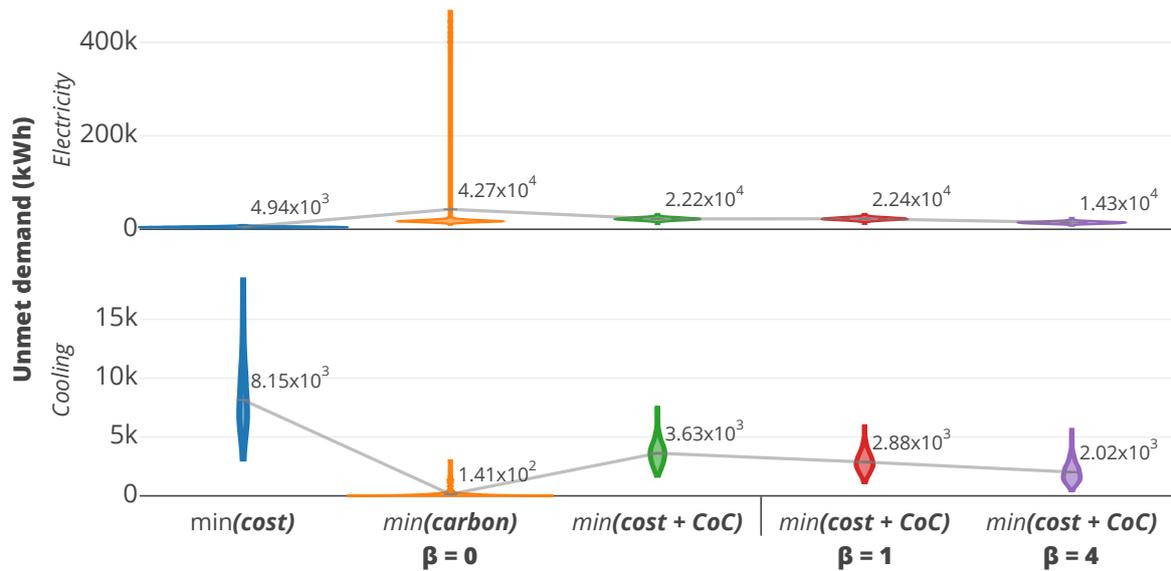


Figure 6.21 Distribution of unmet cooling and electricity demand realised when running OOS tests on the result of 10-scenario risk-neutral SO and risk-averse SO in the Bangalore case study district. Risk-averse models consider the cost of carbon in the CVaR component of the objective function. CoC = cost of carbon,  $\beta$  = risk aversion measure. Values adjacent to distributions refer to their mean.

## 6.3 Discussion

It is clearly difficult to justify the monetary expense involved in implementing centralised energy systems. But, such systems do exist and are regularly considered by district master-planners. Understanding the impact of decision-maker constraints is therefore important. For instance, investment limits may prove beneficial, if those limits exceed the investment cost given by the balanced optimal solution. In the UK context, this would allow a marginally increased prevalence of PV installations. Only by applying constraints to enforce the construction of a centralised system are they installed in favour of building-level technologies in the UK. Enforcement may mean a carbon objective, a cost of carbon, or a strict limit on building-level energy provision. If any enforcements are implemented, investment costs will increase by at least an order of magnitude. In the Bangalore context, centralised cooling provision becomes viable for a smaller increase in investment cost, presenting a lower barrier to a decision maker being influenced by auxiliary incentives, such as the minimisation of carbon emissions.

A cost of carbon may be a feasible mechanism to incentivise decision makers to move towards technologies with lower levels of CO<sub>2</sub> emissions. Yet, more than half of carbon emission savings would not be realised using the current expected best-case cost of carbon in 2050. In Bangalore, if the ‘politically’ viable cost of carbon of 5 USD/tCO<sub>2</sub> (Hourcade *et al.*, 2017) were to be applied, it would not be financially worthwhile to make any changes to the cost-optimal system. As such, it is unlikely that a cost of carbon will affect district level decision makers.

A primary reason to push for district systems is their perceived lower emissions. Indeed, the results in this chapter agree: the lowest carbon systems are certainly those which make most use of centralised energy generation and/or storage. But, this is due to the lack of a

barrier in the model associated with technology investment. A greater number of technologies are required to be linked together to meet the lowest carbon emission objective. This raises a number of concerns not addressed in this thesis. First, maintenance requirements will be necessarily greater. Second, if the embodied carbon of technologies could be reliably calculated, it is likely that they would not only increase linearly with technology capacity, but would also have some fixed value associated with the existence of a technology, no matter the size. Third, the need to improve control systems increases. Additionally, in Bangalore, the use of a [DG](#) increases markedly at low carbon emission levels, but this ignores the particulate emissions which would inevitably impact health of those in the immediate vicinity. The same can be said for a biomass fired [CCHP](#) (Omu *et al.*, 2015).

If a centralised system is to be installed, it is also important to understand whether it is worth connecting the whole district. Results from the Cambridge case point to either ignoring some buildings, or finding ways to reach those buildings which does not require the network to follow an existing road network. Around a third of the buildings resist full connection to the grid as the district heat demand is better met by centralised generation. Thus, it is unlikely to be worthwhile to extend the network throughout the district, particularly as the carbon emissions do not markedly reduce.

Increasing centralised energy provision does provide the benefit of decreasing the unmet demand in the Bangalore district, when running [OOS](#) tests. The greater flexibility provided by the installed [StoreT](#) leads to a more robust system for meeting cooling demand. But, there is no marked benefit in the amount of electricity demand that is met. The latter is also probably the most important demand to be met, corresponding to a complete inability to conduct office work, while unmet cooling demand corresponds only to occupant discomfort. Of course, occupant discomfort can also lead to the inability to conduct office work, but it is more gradual.

Increasing the allowable system investment cost does not automatically confer an improvement in system robustness to unmet demand. But, it seems from the results of both sets of  $\epsilon$ -constrained investment cost [OOS](#) test results that any investment portfolio is better than the balanced cost-optimal option. The baseline cost model has much higher unmet electricity demand when minimising carbon and higher unmet cooling demand when minimising cost, compared to systems with both lower and higher associated investment cost. From an unmet electricity demand perspective, the completely conventional system achieved by investment cost minimisation would be ideal. Certainly, lower investment cost systems which do not inherently account for demand fare better when unexpected demand is met.

Designing a system to minimise carbon would be the least helpful in ensuring system robustness, as [OOS](#) tests indicate a high expected unmet electricity demand relative to all other designs. Moreover, there are some worst-case [OOS](#) scenarios which would lead to the system failing, causing unmet demand orders of magnitude greater than the mean. These outlier scenarios are unique to this particular objective. An alternative compromise may be to account for carbon emissions in the cost minimisation objective. This would require a cost of carbon greater than the politically viable cost for India, by at least an order of magnitude. Nevertheless, if put in place, a 9.49 [INR](#)/kgCO<sub>2</sub> cost of carbon would lead to a

system which is more robust to unmet demand than that achieved by optimising either cost or carbon individually, whilst also reducing carbon emissions compared to the cost-optimal design. Minimising the cost of carbon alongside other monetary costs would result in the use of a centralised biomass fuelled CCHP to meet most cooling demand, but would still require heavy dependence on the national grid for electricity. Of course, such dependence on grid electricity would cause significant system reliability issues if allowing electricity interruptions to be considered, as done previously in Section 5.5.2. Only carbon emission minimisation leads to a system independent of the national electricity network. Converging on an investment which adds electrical grid interruptions to the objective function is a natural extension to this study which would lead to a system which balances cost, carbon, demand uncertainty, and supply instability.

## 6.4 Conclusions

Investment decisions do change when different objectives are considered for the same case study district. In this chapter the traditional objective of a minimised balanced monetary cost has been compared to one where varying importance is applied to investment and operation monetary cost, carbon emissions and monetary costs, and dependence on the centralised generation of thermal energy. The minimisation of balanced monetary cost inevitably leads to the selection of conventional, building-level technologies. Introducing new objectives can change this, although the impact may only be small.

In Cambridge, the combination of an [natural gas boiler \(NGB\)](#) to meet heat demand and [national grid electricity \(GridE\)](#) to meet electricity demand is not only commonplace in buildings today, but is optimal from a cost perspective. Increasing the limit on investment costs, or reducing the limit on carbon emissions, initially serves to increase [solar photovoltaic panel \(PV\)](#) capacity. Subsequent investment in a [solar thermal panel \(ST\)](#) and [thermal energy storage \(StoreT\)](#) occurs before centralised energy generation can be deemed to be viable. Once investment in an energy centre has occurred, investment costs will have increased by several orders of magnitude. A 10% reduction in carbon emissions could be realised by the introduction of a 0.07 [GBP/kgCO<sub>2</sub>](#) cost of carbon, but would require the introduction of costs currently at the upper extreme of those expected to be imposed.

In Bangalore, the cost-optimal solution is similar: an [electric chiller \(ECh\)](#) to meet cooling demand is coupled with a [GridE](#) connection to meet electricity demand. Unlike the Cambridge case, due to the differences in costs and climate, [PV](#) is also central to a cost-optimal solution in Bangalore. However, this dependence on [PV](#) rapidly dwindles if investment costs are increasingly constrained. The district system is cheaper to implement than in Cambridge, leading to the use of a centralised energy centre in many more situations. Due to the current emissions factor associated with dependence on [GridE](#), building-level [diesel generators \(DGs\)](#) are the carbon-optimal method to meet electricity demand. Similar to the Cambridge case study, a 10% reduction in carbon emissions could be realised within the range of carbon costs that could feasibly be imposed. But, a much lower ‘politically’ viable cost of carbon would have no impact on the system design.

There are variations in the ability of different investment decisions to meet unexpected demand scenarios. Most of these variations are unintended, as none of the objectives are directed towards mitigating demand uncertainty. Unmet cooling demand can be reduced if using a centralised energy centre. However, this can be coupled with an increase in unmet electricity demand, probably a more important energy carrier to ensure that is met in office buildings. Both electricity and unmet demand are reduced by the incorporation of [scenario optimisation \(SO\)](#), a conclusion also drawn in chapter 5. The incorporation of carbon emissions reduction in the [SO](#) objective produces a cost-carbon-risk minimisation function which is promising as a method of increasing system robustness to unmet demand, excessive carbon emissions and excessive operational costs.



# Chapter 7

## Conclusions

This research aimed to improve methods of district energy system optimisation to make it practical as a tool in the decision-making process. This has been achieved by the completion of three methodological enhancements and the development of a modelling framework. The development of the Calliope framework is central to the work undertaken throughout this thesis. It is a [mixed integer linear programming \(MILP\)](#) framework, openly available under the Apache 2.0 licence, which offers powerful processes to prepare a model for optimisation and to analyse the results. Calliope can be used to answer questions about optimal technology investment capacity and operation for energy systems in which any number of locations, energy carriers, and time scales exist. An up-to-date Python toolchain ensures that models can be defined in an easy to read manner, data can be processed efficiently and shared seamlessly, and visualisations can be interactive. Moreover, mature [MILP](#) algorithms are available within the framework, either through commercial solvers such as IBM ILOG CPLEX or openly available solvers such as GLPK. Calliope was originally developed to address questions on national-scale energy infrastructure and has been extended in this research to enable district-scale energy systems to be effectively modelled. Additionally, its internal and external structure was updated to increase its accessibility to users outside the academic community.

A review of the literature which encompasses the current state-of-the-art in district energy system optimisation highlighted shortcomings which limit the viability of using optimisation in practical district design. Each methodological enhancement detailed in this thesis was aimed at addressing these shortcomings. First, [Chapter 4](#) examined how technologies are represented in models. Most importantly, the impact of technology part-load consumption curve representation on solution time, optimal operation schedule and capacity investment were tested. The metaheuristic approach to an energy system model may offer a more accurate representation of reality, but time is required to reach a solution. And, even though more accurate, it may not be globally optimal. Piecewise linearisation was demonstrated as an effective middle ground between the nonlinear technology part-load curves and the use of a rated efficiency. Instead of needing a nonlinear representation of a characteristic, a combination of linear pieces are described along a piecewise curve. It offers a reasonable fit to nonlinear data whilst allowing a much quicker, [MILP](#) model to be implemented. This

comparison has not previously been quantified in energy system optimisation. In fact, the impact of simplification from metaheuristic to **MILP** models is too often ignored. Further investigation showed that advancements on traditional piecewise linearisation are possible, improving both the accuracy of results and the solution time of **MILP** models. By automating the process, it is possible for a decision-maker to only supply nonlinear curves to Calliope. Conversion of nonlinearities to viable piecewise curves would be undertaken by a pre-processing step, thus preparing the curves for application in **MILP** modelling. However, an increase in solution time should be expected. This cannot be avoided and may be exacerbated by the particular model being solved. If the possibility of part-load operation is near-optimal, convergence on a solution can take time. Indeed, this concern led to the stand-alone analysis of chapter 4; part-load characteristics were not extended to chapters 5 or 6.

Chapter 5 considered the incorporation of uncertainty in linear modelling. Using a three-step method to generate, reduce, and optimise scenarios of building demand, more robust investment portfolios were achieved. Unlike previous research in the field of **scenario optimisation (SO)**, this research used historical demand data to produce multidimensional probability density functions, which were sampled to generate demand scenarios. Additionally, the choice of method to select scenarios, namely, by reducing the probability distance between independently optimal **objective function values (OFVs)**, is an enhancement on existing approaches for district energy scenario optimisation. It thus proved straightforward to prepare a 10 scenario **SO** model from 500 generated demand scenarios. The result of the three-step method was compared to a single scenario model representing the typical approach to energy systems modelling based on mean demand data. **SO** increased the required technology capacity in both case studies, particularly for conventional energy technologies such as building-level **natural gas boilers (NGBs)** and **electric chillers (EChs)**. In Bangalore, mean model **solar photovoltaic panel (PV)** capacity was reduced when accounting for uncertainty, in favour of greater conventional technology capacity. While this reduced investment costs, it increased the expected operational costs. In Cambridge, storage capacity was also introduced, to provide greater flexibility for an increase in investment cost. The investment portfolio from both case studies proved to be significantly more robust to unmet demand than the mean model. This was tested using **out-of-sample (OOS)** scenarios, which operated on a rolling horizon to expose the fixed technology capacities to each of the 500 demand scenarios in parallel. A two order of magnitude reduction in unmet thermal and electricity demand was achieved using **SO** results compared to those from the mean models. The investment portfolios from the independently optimal models with highest **OFV** and total annual demand were also tested, achieving improvements on the mean models. However, this still leads to an order of magnitude greater unmet demand than the **SO** results.

Enhancements to **SO** were also explored. Particularly, the **conditional value at risk (CVaR)** measure was applied to models to help mitigate the worst-case realisation of costs. This served to further reduce unmet demand and the highest expected cost that could be incurred in the most extreme scenario(s). But it also caused the overall expected cost of the system to increase. The **CVaR** measure was further adapted to consider the dependence of the Bangalore case study district on the **national grid electricity (GridE)**. Unlike the Cambridge

case study, Bangalore districts will be expected to withstand regular interruptions to power. Low system resilience was shown for the mean and standard **SO** models when exposing them to unexpected power interruptions as part of rolling horizon **OOS** tests. Penalising the dependence on the **GridE** led to the district depending heavily on **diesel generators (DGs)** for electricity provision. It also led to some use of the previously unused district-scale centralised network, through investment in some **biomass fuelled CCHP (B-CCHP)** capacity. This led to a marked reduction in unmet demand when both demand scenarios and power interruptions were included in the **OOS** tests. But, an approximately 50% increase in investment cost is incurred to ensure this robustness. Intermittent **GridE** was also added to the mean model, by setting its availability to zero in particular hours of the day. This was not as effective at reducing unmet demand as the **SO** method, but was similarly costly. Finally, carbon emissions were added to the **CVaR** measure in Chapter 6. Combining monetary cost, carbon emissions, and demand uncertainty in the objective function led to the design of a system which aimed to balance each component. In this system, there is an increased dependence on storage technologies and an equal increase in **DG** and **B-CCHP** capacity. The capacity of these new technologies is small, but leads to an expected reduction in carbon emissions of 10% for a corresponding increase in monetary cost of 10%. Using carbon emissions as the risk measure reduces the standard deviation of emissions across the 10 scenarios used in **SO**, but increases the standard deviation of monetary cost: it is not possible to completely balance the two competing components. The unmet demand from **OOS** tests is balanced, however. It occupies a lower range of unmet electricity demand and unmet cooling demand when compared to the result of minimising only for carbon emissions and monetary cost, respectively.

As well as updating the **SO** objective function, Chapter 6 investigates the impact of objective on the Cambridge and Bangalore mean models. The typical single scenario objective function, which balances the annualised investment cost with operational costs, does not necessarily reflect the decision-making process; different decision makers may have different objectives. Therefore, eight different objectives are modelled, including: the cost-focussed developer or occupant, who prefers to minimise one of investment or operation cost; the carbon-constrained decision maker, who may want to minimise cost, but is obliged to meet carbon emission reduction targets, and; the district-focussed client, whose aim is to implement a district system, irrespective of the monetary cost involved. As expected, investment decisions change as each decision maker is considered and consequently there are large variations in possible system costs. Particularly in the Cambridge case study, energy provision centralised at the district level is not optimal unless it is forced by a constraint which requires it or by the need to have extremely low carbon emissions. In either case, investment cost increases by several orders of magnitude and is unlikely to be overcome. Compromises could be achieved by investment in **PV** and **solar thermal panel (ST)** technologies, to at least reduce carbon emissions. In Bangalore, a relatively small increase in the limit on investment cost would begin to see a system centralised at the district level being deployed, which could make it viable to introduce district cooling networks in the city.

Each of the methodological steps taken in this research were aimed at practical optimisation of district energy systems. This has led to further understanding of the impact of model simplification, demand uncertainty, and objective. Through implementation of this updated Calliope framework, a decision maker can be confident in their understanding of model results, and the limitations by which they are constrained. Furthermore, extensible and transparent processes have been implemented to allow complex system data to be applied to a model. A modeller need not decide the best way to piecewise linearise a nonlinear characteristic curve or to generate scenarios from stochastic datasets. Instead, the workflow developed in this research is capable of applying state-of-the-art methods to pre-process complex data in preparation for linear optimisation. This includes breakpoint allocation to minimise linearisation error and allow for the use of [convex bounding sets \(CBS\)](#); scenario generation from historical data using [kernel density estimation \(KDE\)](#) and 5-fold cross-validation, if necessary; and scenario reduction using the fast-forward method applied to independently optimal [OFVs](#). Ex-ante and ex-post error quantification methods are then available to understand the limitations of simplifying the input data, but to also quantify the benefit of using more complex methods than those typical to linear district energy optimisation models. This has been achieved alongside extensibility and transparency. As an openly available, easy to use framework, Calliope can be further developed to meet bespoke decision-maker requirements without losing any of the functionality developed prior to, or during this research.

## 7.1 Limitations and further work

[MILP](#) models are consistently limited by their need to simplify parameters and constraints to ensure a model is tractable. This thesis has introduced a number of methods to simplify intelligently and to better understand the impact of those simplifications. However, further work is still required. [MILP](#) models can be best simplified in two ways: reducing the size of dimensions and reducing the number of binary/integer variables. The time dimension is often reduced first. Indeed, the most aggressive simplification made in this research was timeseries aggregation. Spatial aggregation is understudied, however, indicating a clear avenue of further research into the most effective methods to reduce the number of demand nodes without losing understanding of the district network. Although fewer technologies would also simplify the model, the purpose of optimising district energy systems is to decide between a wide variety of different, often interconnected technologies. Therefore, they cannot be removed from the model. Their representation could be simplified, however, by removing more binary and integer variables. In this thesis, the use of [CBS](#) instead of [special order set constraint of type 2 \(SOS2\)](#) in piecewise linearisation was one method given to reduce the number of binary variables in the model. Technology purchase binary variables were not considered for removal, even though some models failed to converge as a result of their inclusion. They are deemed a necessary component in district energy system optimisation since they place a fixed charge on the installation of technologies, irrespective of their capacity. This emulates the construction of an energy centre, ground work for pipelines, or the cost of

using some building floor area for services. Circumventing their use requires further research, including the use of variable relaxation techniques and multi-stage MILP modelling.

A concern raised from analysing the literature at the beginning of this thesis was that of model data validity. Key to improving on the inability to validate data has been the transparency of this research, leading to public availability of case studies 3 (Cambridge)<sup>1</sup> and 4 (Bangalore)<sup>2</sup>. Yet, there is more to be done. Technology characteristic curves should be validated using operational data of existing facilities. The direct use of such data would ensure that the reality of technology operation is inherent in the model, including the complex interaction of characteristics which may otherwise be ignored. Technology costs should be validated against as-built costs of district systems. This research improved on existing studies by using an industry standard reference for costs in the UK case studies (AECOM, 2015). But, this still ignores the uncertainty in realised costs and is limited to UK studies.

Further work on data validation would inevitably lead to an increase in stochastic input parameters. This research concentrated on demand and power interruption uncertainty. But, it is clear that every parameter can be described by a probability density function (PDF) of some kind. For most, it is unlikely that datasets exist from which PDFs can be developed following the methods used in this thesis. Thus, synonymous with validation concerns, further data collection is required. Additionally, to incorporate several sources of uncertainty into a single model requires more sophisticated scenario selection methods. The number of conceivable, unique scenarios increases considerably with the introduction of new uncertain parameters. Thus, scenario generation and reduction techniques must be developed to best handle greater volumes of data, much of which will be interconnected.

Finally, a near-term goal following this research is the inclusion of methods into the core Calliope offering. The process of introducing new functionality into the stable version of any software requires robust testing, complete documentation, and updates to the user interface that make its use clear and simple. In parallel with the stable Calliope software, piecewise linearisation and scenario optimisation techniques have been developed. They have been developed in the open and are usable for anyone wishing to validate the results of this research or use them for their own analysis. However, the user-facing functionality requires further development to ensure the process is generalised for application to energy systems of all scales before incorporation into a stable release. The inclusion of this functionality will position Calliope as a practical tool for decision makers and will provide the foundation for further development of methods within the field of energy system optimisation.

---

<sup>1</sup><https://github.com/brynpickering/cambridge-calliope>

<sup>2</sup><https://github.com/brynpickering/bangalore-calliope>



# References

- AECOM, ed. (Oct. 2015). *Spon's Mechanical and Electrical Services Price Book 2016*. English. 47 Har/Psc edition. CRC Press. ISBN: 978-1-4987-3506-3.
- Ahn, Y. and Han, J. (June 2018). 'Economic Optimization of Integrated Network for Utility Supply and Carbon Dioxide Mitigation with Multi-Site and Multi-Period Demand Uncertainties'. In: *Applied Energy* 220, pp. 723–734. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2018.02.129](https://doi.org/10.1016/j.apenergy.2018.02.129).
- Akbar, S. *et al.* (2014). *Climate-Smart Development: Adding up the Benefits of Actions That Help Build Prosperity, End Poverty and Combat Climate Change*. Working Paper 88908. Washington, DC: World Bank Group. URL: <http://documents.worldbank.org/curated/en/794281468155721244/Main-report> (visited on 20/09/2018).
- Akbari, K., Jolai, F. and Ghaderi, S. (2016). 'Optimal Design of Distributed Energy System in a Neighborhood under Uncertainty'. In: *Energy* 116, pp. 567–582. DOI: [10.1016/j.energy.2016.09.083](https://doi.org/10.1016/j.energy.2016.09.083).
- Akbari, K. *et al.* (2014). 'Optimal Investment and Unit Sizing of Distributed Energy Systems under Uncertainty: A Robust Optimization Approach'. In: *Energy and Buildings* 85, pp. 275–286. DOI: [10.1016/j.enbuild.2014.09.009](https://doi.org/10.1016/j.enbuild.2014.09.009).
- Al Rafea, K. *et al.* (Mar. 2017). 'Integration of Decentralized Energy Systems with Utility-Scale Energy Storage through Underground Hydrogen Natural Gas Co-Storage Using the Energy Hub Approach'. In: *Industrial & Engineering Chemistry Research* 56.8. WOS:000395493600039, pp. 2310–2330. ISSN: 0888-5885. DOI: [10.1021/acs.iecr.6b02861](https://doi.org/10.1021/acs.iecr.6b02861).
- Alarcon-Rodriguez, A., Ault, G. and Galloway, S. (June 2010). 'Multi-Objective Planning of Distributed Energy Resources: A Review of the State-of-the-Art'. In: *Renewable and Sustainable Energy Reviews* 14.5, pp. 1353–1366. ISSN: 1364-0321. DOI: [10.1016/j.rser.2010.01.006](https://doi.org/10.1016/j.rser.2010.01.006).
- Alderson, D. L. and Doyle, J. C. (July 2010). 'Contrasting Views of Complexity and Their Implications For Network-Centric Infrastructures'. In: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 40.4, pp. 839–852. ISSN: 1083-4427. DOI: [10.1109/TSMCA.2010.2048027](https://doi.org/10.1109/TSMCA.2010.2048027).
- Allegrini, J. *et al.* (Dec. 2015). 'A Review of Modelling Approaches and Tools for the Simulation of District-Scale Energy Systems'. In: *Renewable and Sustainable Energy Reviews* 52, pp. 1391–1404. ISSN: 1364-0321. DOI: [10.1016/j.rser.2015.07.123](https://doi.org/10.1016/j.rser.2015.07.123).
- Ameri, M. and Besharati, Z. (Jan. 2016). 'Optimal Design and Operation of District Heating and Cooling Networks with CCHP Systems in a Residential Complex'. In: *Energy and Buildings* 110, pp. 135–148. ISSN: 0378-7788. DOI: [10.1016/j.enbuild.2015.10.050](https://doi.org/10.1016/j.enbuild.2015.10.050).
- Babonneau, F., Caramanis, M. and Haurie, A. (2017). 'ETEM-SG: Optimizing Regional Smart Energy System with Power Distribution Constraints and Options'. In: *Environmental Modeling and Assessment* 22.5, pp. 411–430. DOI: [10.1007/s10666-016-9544-0](https://doi.org/10.1007/s10666-016-9544-0).
- Baetens, R. and Saelens, D. (2016). 'Modelling Uncertainty in District Energy Simulations by Stochastic Residential Occupant Behaviour'. In: *Journal of Building Performance Simulation* 9.4, pp. 431–447. DOI: [10.1080/19401493.2015.1070203](https://doi.org/10.1080/19401493.2015.1070203).
- BESCOM – Bangalore Electricity Supply Company Ltd. (2017). *15th Annual Report*. Financial Report BESCOM/BC-4/CS/AGM/2017-18/F-22/1044-1066. Bangalore, India. URL: [https://bescom.org/wp-content/uploads/2018/05/Bescom-AnnRep\\_2016-17COLOUR.compressed.pdf](https://bescom.org/wp-content/uploads/2018/05/Bescom-AnnRep_2016-17COLOUR.compressed.pdf) (visited on 13/12/2018).

- Basu, A. K. (Jan. 2013). 'Microgrids: Planning of Fuel Energy Management by Strategic Deployment of CHP-Based DERs – An Evolutionary Algorithm Approach'. In: *International Journal of Electrical Power & Energy Systems* 44.1, pp. 326–336. ISSN: 0142-0615. DOI: [10.1016/j.ijepes.2012.07.059](https://doi.org/10.1016/j.ijepes.2012.07.059).
- Beale, E. M. L. and Tomlin, J. A. (1970). 'Special Facilities in a General Mathematical Programming System for Non-Convex Problems Using Ordered Sets of Variables'. In: *OR* 69.447-454, p. 99.
- Bergstra, J., Yamins, D. and Cox, D. D. (2013). 'Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures'. en. In: *Proceedings of the 30th International Conference on Machine Learning*. Vol. 28. Atlanta, Georgia, USA: JMLR: W&CP, p. 9.
- Best, R. E., Flager, F. and Lepech, M. D. (Dec. 2015). 'Modeling and Optimization of Building Mix and Energy Supply Technology for Urban Districts'. In: *Applied Energy* 159, pp. 161–177. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2015.08.076](https://doi.org/10.1016/j.apenergy.2015.08.076).
- Bischi, A. *et al.* (Sept. 2014). 'A Detailed MILP Optimization Model for Combined Cooling, Heat and Power System Operation Planning'. en. In: *Energy* 74, pp. 12–26. ISSN: 03605442. DOI: [10.1016/j.energy.2014.02.042](https://doi.org/10.1016/j.energy.2014.02.042).
- Braun, M. R., Altan, H. and Beck, S. B. M. (Oct. 2014). 'Using Regression Analysis to Predict the Future Energy Consumption of a Supermarket in the UK'. In: *Applied Energy* 130, pp. 305–313. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2014.05.062](https://doi.org/10.1016/j.apenergy.2014.05.062).
- Bruninx, K. and Delarue, E. (Apr. 2016). 'Scenario Reduction Techniques and Solution Stability for Stochastic Unit Commitment Problems'. In: *2016 IEEE International Energy Conference (ENERGYCON)*, pp. 1–7. DOI: [10.1109/ENERGYCON.2016.7514074](https://doi.org/10.1109/ENERGYCON.2016.7514074).
- Brunold, S., Frey, R. and Frei, U. (Dec. 1994). 'A Comparison of Three Different Collectors for Process Heat Applications'. English. In: URL: <https://www.osti.gov/scitech/biblio/194814> (visited on 27/06/2017).
- Bucciarelli, M., Paoletti, S. and Vicino, A. (Sept. 2018). 'Optimal Sizing of Energy Storage Systems under Uncertain Demand and Generation'. In: *Applied Energy* 225, pp. 611–621. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2018.03.153](https://doi.org/10.1016/j.apenergy.2018.03.153).
- Bukhsh, W. A., Papakonstantinou, A. and Pinson, P. (Apr. 2016). 'A Robust Optimisation Approach Using CVaR for Unit Commitment in a Market with Probabilistic Offers'. In: *2016 IEEE International Energy Conference (ENERGYCON)*, pp. 1–6. DOI: [10.1109/ENERGYCON.2016.7514076](https://doi.org/10.1109/ENERGYCON.2016.7514076).
- Buoro, D. *et al.* (Sept. 2013). 'Multicriteria Optimization of a Distributed Energy Supply System for an Industrial Area'. In: *Energy* 58, pp. 128–137. ISSN: 0360-5442. DOI: [10.1016/j.energy.2012.12.003](https://doi.org/10.1016/j.energy.2012.12.003).
- Buoro, D., Pinamonti, P. and Reini, M. (July 2014). 'Optimization of a Distributed Cogeneration System with Solar District Heating'. In: *Applied Energy* 124, pp. 298–308. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2014.02.062](https://doi.org/10.1016/j.apenergy.2014.02.062).
- Cai, Y. P. *et al.* (May 2009). 'Community-Scale Renewable Energy Systems Planning under Uncertainty—An Interval Chance-Constrained Programming Approach'. In: *Renewable and Sustainable Energy Reviews* 13.4, pp. 721–735. ISSN: 1364-0321. DOI: [10.1016/j.rser.2008.01.008](https://doi.org/10.1016/j.rser.2008.01.008).
- Cambridge City Council (June 2016). *Cambridge West Site*. Planning Application 16/1134/OUT. Cambridge City Council. URL: <https://idox.cambridge.gov.uk/online-applications> (visited on 09/08/2018).
- Campos Celador, A., Odriozola, M. and Sala, J. M. (Aug. 2011). 'Implications of the Modelling of Stratified Hot Water Storage Tanks in the Simulation of CHP Plants'. In: *Energy Conversion and Management* 52.8–9, pp. 3018–3026. ISSN: 0196-8904. DOI: [10.1016/j.enconman.2011.04.015](https://doi.org/10.1016/j.enconman.2011.04.015).
- Cao, K.-K. *et al.* (Sept. 2016). 'Raising Awareness in Model-Based Energy Scenario Studies—A Transparency Checklist'. In: *Energy, Sustainability and Society* 6.1, p. 28. ISSN: 2192-0567. DOI: [10.1186/s13705-016-0090-z](https://doi.org/10.1186/s13705-016-0090-z).
- Capuder, T. and Mancarella, P. (July 2014). 'Techno-Economic and Environmental Modelling and Optimization of Flexible Distributed Multi-Generation Options'. In: *Energy* 71, pp. 516–533. ISSN: 0360-5442. DOI: [10.1016/j.energy.2014.04.097](https://doi.org/10.1016/j.energy.2014.04.097).

- Carpaneto, E., Lazzeroni, P. and Repetto, M. (Mar. 2015). 'Optimal Integration of Solar Energy in a District Heating Network'. In: *Renewable Energy* 75.Supplement C, pp. 714–721. ISSN: 0960-1481. DOI: [10.1016/j.renene.2014.10.055](https://doi.org/10.1016/j.renene.2014.10.055).
- Casasi, M., Pinamonti, P. and Reini, M. (Dec. 2009). 'Optimal Lay-out and Operation of Combined Heat & Power (CHP) Distributed Generation Systems'. In: *Energy*. ECOS 2007 34.12, pp. 2175–2183. ISSN: 0360-5442. DOI: [10.1016/j.energy.2008.10.019](https://doi.org/10.1016/j.energy.2008.10.019).
- cBalance – cBalance Solutions (2012). *GHG Inventory Report for Electricity Generation and Consumption in India*. White Paper. cBalance Solutions Pvt. Ltd. URL: [http://cbalance.in/wp-content/uploads/2013/01/cbalance\\_white-paper\\_Electricity-emission-factors\\_28Dec2012\\_revised\\_V21.pdf](http://cbalance.in/wp-content/uploads/2013/01/cbalance_white-paper_Electricity-emission-factors_28Dec2012_revised_V21.pdf) (visited on 06/11/2018).
- Ceseña, E. A. M., Capuder, T. and Mancarella, P. (Jan. 2016). 'Flexible Distributed Multienergy Generation System Expansion Planning Under Uncertainty'. In: *IEEE Transactions on Smart Grid* 7.1, pp. 348–357. ISSN: 1949-3053. DOI: [10.1109/TSG.2015.2411392](https://doi.org/10.1109/TSG.2015.2411392).
- Cesena, E. and Mancarella, P. (2018). 'Energy Systems Integration in Smart Districts: Robust Optimisation of Multi-Energy Flows in Integrated Electricity, Heat and Gas Networks'. Article in Press. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85045726618%5C&doi=10.1109%5C%2fTSG.2018.2828146%5C&partnerID=40%5C&md5=6ac0fb7295103d98ec53953eccfa91ae>.
- Chen, Y. *et al.* (2016). 'Regional Planning of New-Energy Systems within Multi-Period and Multi-Option Contexts: A Case Study of Fengtai, Beijing, China'. In: *Renewable and Sustainable Energy Reviews* 65, pp. 356–372. DOI: [10.1016/j.rser.2016.07.017](https://doi.org/10.1016/j.rser.2016.07.017).
- Chiang, M. M.-T. and Mirkin, B. (Mar. 2010). 'Intelligent Choice of the Number of Clusters in K-Means Clustering: An Experimental Study with Different Cluster Spreads'. *en*. In: *Journal of Classification* 27.1, pp. 3–40. ISSN: 0176-4268, 1432-1343. DOI: [10.1007/s00357-010-9049-5](https://doi.org/10.1007/s00357-010-9049-5).
- Chicco, G. and Mancarella, P. (Apr. 2009). 'Distributed Multi-Generation: A Comprehensive View'. In: *Renewable and Sustainable Energy Reviews* 13.3, pp. 535–551. ISSN: 1364-0321. DOI: [10.1016/j.rser.2007.11.014](https://doi.org/10.1016/j.rser.2007.11.014).
- Christidis, A. *et al.* (May 2012). 'The Contribution of Heat Storage to the Profitable Operation of Combined Heat and Power Plants in Liberalized Electricity Markets'. In: *Energy*. 23rd International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems, ECOS 2010 41.1, pp. 75–82. ISSN: 0360-5442. DOI: [10.1016/j.energy.2011.06.048](https://doi.org/10.1016/j.energy.2011.06.048).
- Conejo, A. J., Carrión, M. and Morales, J. M. (2010). *Decision Making under Uncertainty in Electricity Markets*. Vol. 1. Springer.
- Connolly, D. *et al.* (Feb. 2014). 'Heat Roadmap Europe: Combining District Heating with Heat Savings to Decarbonise the EU Energy System'. In: *Energy Policy* 65, pp. 475–489. ISSN: 0301-4215. DOI: [10.1016/j.enpol.2013.10.035](https://doi.org/10.1016/j.enpol.2013.10.035).
- D'Ambrosio, C., Lodi, A. and Martello, S. (Jan. 2010). 'Piecewise Linear Approximation of Functions of Two Variables in MILP Models'. In: *Operations Research Letters* 38.1, pp. 39–46. ISSN: 0167-6377. DOI: [10.1016/j.orl.2009.09.005](https://doi.org/10.1016/j.orl.2009.09.005).
- De Césaró Oliveski, R., Krenzinger, A. and Vielmo, H. A. (Aug. 2003). 'Comparison between Models for the Simulation of Hot Water Storage Tanks'. In: *Solar Energy* 75.2, pp. 121–134. ISSN: 0038-092X. DOI: [10.1016/j.solener.2003.07.009](https://doi.org/10.1016/j.solener.2003.07.009).
- DeCarolis, J. F., Hunter, K. and Sreepathi, S. (Nov. 2012). 'The Case for Repeatable Analysis with Energy Economy Optimization Models'. In: *Energy Economics* 34.6, pp. 1845–1853. ISSN: 0140-9883. DOI: [10.1016/j.eneco.2012.07.004](https://doi.org/10.1016/j.eneco.2012.07.004).
- del Valle, Y., Harley, R. G. and Venayagamoorthy, G. K. (2009). 'Comparison of Enhanced-PSO and Classical Optimization Methods: A Case Study for STATCOM Placement'. In: *Intelligent System Applications to Power Systems, 2009. ISAP'09. 15th International Conference On*. IEEE, pp. 1–7. URL: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=5352884](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5352884) (visited on 15/01/2016).

- BEIS – Department for Business, Energy & Industrial Strategy (July 2018a). ‘Digest of United Kingdom Energy Statistics (DUKES) 2018’. In: itemType: dataset. URL: <https://www.gov.uk/government/statistics/digest-of-uk-energy-statistics-dukes-2018-main-report> (visited on 17/12/2018).
- Díaz Redondo, P. and van Vliet, O. (Aug. 2015). ‘Modelling the Energy Future of Switzerland after the Phase Out of Nuclear Power Plants’. In: *Energy Procedia*. European Geosciences Union General Assembly 2015 - Division Energy, Resources and Environment, EGU 2015 76, pp. 49–58. ISSN: 1876-6102. DOI: [10.1016/j.egypro.2015.07.843](https://doi.org/10.1016/j.egypro.2015.07.843).
- Dollar, D., Hallward-Driemeier, M. and Mengistae, T. (Oct. 2005). ‘Investment Climate and Firm Performance in Developing Economies’. In: *Economic Development and Cultural Change* 54.1, pp. 1–31. ISSN: 0013-0079. DOI: [10.1086/431262](https://doi.org/10.1086/431262).
- EQS WS – Energy Quality of Supply Work Stream (July 2018). *CEER Benchmarking Report 6.1 on the Continuity of Electricity and Gas Supply*. Tech. rep. C18-EQS-86-03. Brussels, Belgium: Council of European Energy Regulators. URL: <https://www.ceer.eu/documents/104400/-/-/963153e6-2f42-78eb-22a4-06f1552dd34c> (visited on 29/01/2019).
- Evins, R. *et al.* (Aug. 2014). ‘New Formulations of the ‘Energy Hub’ Model to Address Operational Constraints’. en. In: *Energy* 73, pp. 387–398. ISSN: 03605442. DOI: [10.1016/j.energy.2014.06.029](https://doi.org/10.1016/j.energy.2014.06.029).
- Faille, D., Mondon, C. and Al-Nasrawi, B. (Nov. 2007). ‘mCHP Optimization by Dynamic Programming and Mixed Integer Linear Programming’. In: *International Conference on Intelligent Systems Applications to Power Systems, 2007. ISAP 2007*, pp. 1–6. DOI: [10.1109/ISAP.2007.4441677](https://doi.org/10.1109/ISAP.2007.4441677).
- Farzaneh, H., Doll, C. N. H. and Puppim de Oliveira, J. A. (Feb. 2016). ‘An Integrated Supply-Demand Model for the Optimization of Energy Flow in the Urban System’. In: *Journal of Cleaner Production*. Towards Post Fossil Carbon Societies: Regenerative and Preventative Eco-Industrial Development 114, pp. 269–285. ISSN: 0959-6526. DOI: [10.1016/j.jclepro.2015.05.098](https://doi.org/10.1016/j.jclepro.2015.05.098).
- Fazlollahi, S., Becker, G. and Maréchal, F. (Dec. 2014). ‘Multi-Objectives, Multi-Period Optimization of District Energy Systems: II—Daily Thermal Storage’. In: *Computers & Chemical Engineering* 71. Supplement C, pp. 648–662. ISSN: 0098-1354. DOI: [10.1016/j.compchemeng.2013.10.016](https://doi.org/10.1016/j.compchemeng.2013.10.016).
- Fonseca, J. A. and Schlueter, A. (Mar. 2015). ‘Integrated Model for Characterization of Spatiotemporal Building Energy Consumption Patterns in Neighborhoods and City Districts’. In: *Applied Energy* 142, pp. 247–265. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2014.12.068](https://doi.org/10.1016/j.apenergy.2014.12.068).
- Forrest, J. *et al.* (July 2018). *Coin-or/Cbc: Version 2.9.9*. Zenodo. DOI: [10.5281/zenodo.1317566](https://doi.org/10.5281/zenodo.1317566). URL: <https://zenodo.org/record/1317566> (visited on 01/01/2019).
- Fuentes-Cortés, L. F., Santibañez-Aguilar, J. E. and Ponce-Ortega, J. M. (May 2016). ‘Optimal Design of Residential Cogeneration Systems under Uncertainty’. In: *Computers & Chemical Engineering* 88, pp. 86–102. ISSN: 0098-1354. DOI: [10.1016/j.compchemeng.2016.02.008](https://doi.org/10.1016/j.compchemeng.2016.02.008).
- Gabrielli, P. *et al.* (June 2018). ‘Optimal Design of Multi-Energy Systems with Seasonal Storage’. In: *Applied Energy* 219, pp. 408–424. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2017.07.142](https://doi.org/10.1016/j.apenergy.2017.07.142).
- Gamou, S., Yokoyama, R. and Ito, K. (June 2002). ‘Optimal Unit Sizing of Cogeneration Systems in Consideration of Uncertain Energy Demands as Continuous Random Variables’. In: *Energy Conversion and Management* 43.9, pp. 1349–1361. ISSN: 0196-8904. DOI: [10.1016/S0196-8904\(02\)00020-1](https://doi.org/10.1016/S0196-8904(02)00020-1).
- Garcia, R. S. and Weisser, D. (Nov. 2006). ‘A Wind–Diesel System with Hydrogen Storage: Joint Optimisation of Design and Dispatch’. In: *Renewable Energy* 31.14, pp. 2296–2320. ISSN: 0960-1481. DOI: [10.1016/j.renene.2005.11.003](https://doi.org/10.1016/j.renene.2005.11.003).
- ofgem – Gas and Electricity Markets Authority (Feb. 2016). *Feed-in Tariff Generation & Export Payment Rate Table for Photovoltaic Installations - FIT Year 7 (2016/17)*. Tech. rep. URL: <https://www.ofgem.gov.uk/publications-and-updates/feed-tariff-fit-tariff-table-1-april-2016-pv-only> (visited on 29/03/2019).

- Giraud, L. *et al.* (2017). 'Optimal Control of District Heating Systems Using Dynamic Simulation and Mixed Integer Linear Programming'. In: *Proceedings of the 12th International Modelica Conference, Prague, Czech Republic, May 15-17, 2017*. Linköping University Electronic Press, pp. 141–150.
- GNU Linear Programming Kit, Version 6.1 (2018). URL: <http://www.gnu.org/software/glpk/glpk.html>.
- Goderbauer, S., Comis, M. and Willamowski, F. J. L. (May 2018). 'The Synthesis Problem of Decentralized Energy Systems Is Strongly NP-Hard'. en. In: *repORt (preprint) 2018.043*, p. 16. URL: [http://www.or.rwth-aachen.de/research/publications/Complexity\\_SynthesisEnergySystems\\_GoderbauerComisWillamowski\\_20180516.pdf](http://www.or.rwth-aachen.de/research/publications/Complexity_SynthesisEnergySystems_GoderbauerComisWillamowski_20180516.pdf).
- Gomory, R. E. (1960). *An Algorithm for the Mixed Integer Problem*. en. Product Page. URL: [https://www.rand.org/pubs/research\\_memoranda/RM2597.html](https://www.rand.org/pubs/research_memoranda/RM2597.html) (visited on 26/11/2018).
- Good, N. and Mancarella, P. (2017). 'Flexibility in Multi-Energy Communities with Electrical and Thermal Storage: A Stochastic, Robust Approach for Multi-Service Demand Response'. In: *IEEE Transactions on Smart Grid* PP.99, pp. 1–1. ISSN: 1949-3053. DOI: [10.1109/TSG.2017.2745559](https://doi.org/10.1109/TSG.2017.2745559).
- Grams, C. M. *et al.* (Aug. 2017). 'Balancing Europe's Wind-Power Output through Spatial Deployment Informed by Weather Regimes'. en. In: *Nature Climate Change* 7.8, pp. 557–562. ISSN: 1758-6798. DOI: [10.1038/nclimate3338](https://doi.org/10.1038/nclimate3338).
- GLA – Greater London Authority (Mar. 2016). *The London Plan*. Tech. rep. Mayor of London. URL: [https://www.london.gov.uk/sites/default/files/the\\_london\\_plan\\_2016\\_jan\\_2017\\_fix.pdf](https://www.london.gov.uk/sites/default/files/the_london_plan_2016_jan_2017_fix.pdf) (visited on 13/11/2018).
- Gu, W. *et al.* (Jan. 2014). 'Modeling, Planning and Optimal Energy Management of Combined Cooling, Heating and Power Microgrid: A Review'. In: *International Journal of Electrical Power & Energy Systems* 54, pp. 26–37. ISSN: 0142-0615. DOI: [10.1016/j.ijepes.2013.06.028](https://doi.org/10.1016/j.ijepes.2013.06.028).
- Gurobi – Gurobi Optimization, LLC (2018). *Gurobi Optimizer Reference Manual*. URL: <http://www.gurobi.com>.
- Gustafsson, S.-I., Karlsson, B. G. and Sjöholm, B. H. (Jan. 1987). 'Differential Rates for District Heating and the Influence on the Optimal Retrofit Strategy for Multi-Family Buildings'. In: *Heat Recovery Systems and CHP* 7.4, pp. 337–341. ISSN: 0890-4332. DOI: [10.1016/0890-4332\(87\)90095-0](https://doi.org/10.1016/0890-4332(87)90095-0).
- Haikarainen, C., Pettersson, F. and Saxén, H. (Sept. 2014). 'A Model for Structural and Operational Optimization of Distributed Energy Systems'. In: *Applied Thermal Engineering* 70.1, pp. 211–218. ISSN: 1359-4311. DOI: [10.1016/j.applthermaleng.2014.04.049](https://doi.org/10.1016/j.applthermaleng.2014.04.049).
- Hart, W. E. *et al.* (2017). *Pyomo — Optimization Modeling in Python*. en. 2nd ed. Springer Optimization and Its Applications. Springer International Publishing. ISBN: 978-3-319-58819-3. URL: <http://www.springer.com/de/book/9783319588193> (visited on 29/11/2018).
- Hartigan, J. A. and Hartigan, J. A. (1975). *Clustering Algorithms*. Vol. 209. Wiley New York.
- Hawkes, A. and Leach, M. (July 2009). 'Modelling High Level System Design and Unit Commitment for a Microgrid'. en. In: *Applied Energy* 86.7-8, pp. 1253–1265. ISSN: 03062619. DOI: [10.1016/j.apenergy.2008.09.006](https://doi.org/10.1016/j.apenergy.2008.09.006).
- Hawkey, D., Webb, J. and Winskel, M. (July 2013). 'Organisation and Governance of Urban Energy Systems: District Heating and Cooling in the UK'. In: *Journal of Cleaner Production*. Special Issue: Advancing Sustainable Urban Transformation 50, pp. 22–31. ISSN: 0959-6526. DOI: [10.1016/j.jclepro.2012.11.018](https://doi.org/10.1016/j.jclepro.2012.11.018).
- Henning, D., Amiri, S. and Holmgren, K. (Dec. 2006). 'Modelling and Optimisation of Electricity, Steam and District Heating Production for a Local Swedish Utility'. en. In: *European Journal of Operational Research* 175.2, pp. 1224–1247. ISSN: 03772217. DOI: [10.1016/j.ejor.2005.06.026](https://doi.org/10.1016/j.ejor.2005.06.026).
- Hiremath, R. B., Shikha, S. and Ravindranath, N. H. (June 2007). 'Decentralized Energy Planning; Modeling and Application—a Review'. In: *Renewable and Sustainable Energy Reviews* 11.5, pp. 729–752. ISSN: 1364-0321. DOI: [10.1016/j.rser.2005.07.005](https://doi.org/10.1016/j.rser.2005.07.005).
- Hirst, D. (Jan. 2018). *Carbon Price Floor (CPF) and the Price Support Mechanism*. Briefing Paper 05927. House of Commons.

- Hoke, A. *et al.* (Aug. 2013). 'Look-Ahead Economic Dispatch of Microgrids with Energy Storage, Using Linear Programming'. In: *2013 1st IEEE Conference on Technologies for Sustainability (SusTech)*, pp. 154–161. DOI: [10.1109/SusTech.2013.6617313](https://doi.org/10.1109/SusTech.2013.6617313).
- Hong, T. and Jiang, Y. (Oct. 1995). 'Stochastic Weather Model for Building HVAC Systems'. In: *Building and Environment* 30.4, pp. 521–532. ISSN: 0360-1323. DOI: [10.1016/0360-1323\(95\)00007-S](https://doi.org/10.1016/0360-1323(95)00007-S).
- Horner, A. and Beauchamp, J. (1996). 'Piecewise-Linear Approximation of Additive Synthesis Envelopes: A Comparison of Various Methods'. In: *Computer Music Journal* 20.2, pp. 72–95. ISSN: 0148-9267. DOI: [10.2307/3681333](https://doi.org/10.2307/3681333).
- Hourcade, J.-C. *et al.* (2017). 'How to Use SVMAs to Reduce the Carbon Pricing and Climate Finance Gap: Numerical Illustrations'. In: *How to use SVMAs to reduce the Carbon Pricing and Climate Finance Gap: numerical illustrations*.
- Hoyer, S. and Hamman, J. J. (Apr. 2017). 'Xarray: N-D Labeled Arrays and Datasets in Python'. en. In: *Journal of Open Research Software* 5. ISSN: 2049-9647. DOI: [10.5334/jors.148](https://doi.org/10.5334/jors.148).
- Huang, Z. *et al.* (Feb. 2015). 'Methods and Tools for Community Energy Planning: A Review'. In: *Renewable and Sustainable Energy Reviews* 42, pp. 1335–1348. ISSN: 1364-0321. DOI: [10.1016/j.rser.2014.11.042](https://doi.org/10.1016/j.rser.2014.11.042).
- IBM Corp. (2016). *IBM ILOG CPLEX Optimisation Studio*. Armonk, NY, US.
- IEA Statistics (2014). *Electric Power Transmission and Distribution Losses (% of Output)*. URL: <http://data.worldbank.org/indicator/EG.ELC.LOSS.ZS> (visited on 15/07/2016).
- Ikeda, S. and Ooka, R. (Feb. 2016). 'Optimal Operation of Energy Systems Including Energy Storage Equipment under Different Connections and Electricity Prices'. en. In: *Sustainable Cities and Society* 21, pp. 1–11. ISSN: 22106707. DOI: [10.1016/j.scs.2015.10.007](https://doi.org/10.1016/j.scs.2015.10.007).
- IWG – Interagency Working Group on Social Cost of and Greenhouse Gases (Aug. 2016). *Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis - Under Executive Order 12866*. Tech. rep. United States Government.
- IEA – International Energy Agency (2014). *Linking Heat and Electricity Systems*. Tech. rep. OECD/IEA.
- IPCC (2014). *Climate Change 2014: Mitigation of Climate Change. Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. en. Ed. by O. Edenhofer *et al.* IPCC 5. OCLC: 892580682. New York, NY: Cambridge Univ. Press. ISBN: 978-1-107-05821-7.
- Jain, R. K. *et al.* (June 2014). 'Forecasting Energy Consumption of Multi-Family Residential Buildings Using Support Vector Regression: Investigating the Impact of Temporal and Spatial Monitoring Granularity on Performance Accuracy'. In: *Applied Energy* 123, pp. 168–178. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2014.02.057](https://doi.org/10.1016/j.apenergy.2014.02.057).
- Jennings, M., Fisk, D. and Shah, N. (2014). 'Modelling and Optimization of Retrofitting Residential Energy Systems at the Urban Scale'. In: *Energy* 64, pp. 220–233. DOI: [10.1016/j.energy.2013.10.076](https://doi.org/10.1016/j.energy.2013.10.076).
- Jin, M. *et al.* (2018). 'Microgrid to Enable Optimal Distributed Energy Retail and End-User Demand Response'. In: *Applied Energy* 210, pp. 1321–1335. DOI: [10.1016/j.apenergy.2017.05.103](https://doi.org/10.1016/j.apenergy.2017.05.103).
- Kazemzadeh, N., Ryan, S. M. and Hamzeei, M. (Dec. 2017). 'Robust Optimization vs. Stochastic Programming Incorporating Risk Measures for Unit Commitment with Uncertain Variable Renewable Generation'. en. In: *Energy Systems*. ISSN: 1868-3975. DOI: [10.1007/s12667-017-0265-5](https://doi.org/10.1007/s12667-017-0265-5).
- Keirstead, J., Jennings, M. and Sivakumar, A. (Aug. 2012). 'A Review of Urban Energy System Models: Approaches, Challenges and Opportunities'. en. In: *Renewable and Sustainable Energy Reviews* 16.6, pp. 3847–3866. ISSN: 13640321. DOI: [10.1016/j.rser.2012.02.047](https://doi.org/10.1016/j.rser.2012.02.047).
- Keirstead, J., Samsatli, N. and Shah, N. (2010). 'SynCity: An Integrated Tool Kit for Urban Energy Systems Modelling'. en. In: *Fifth Urban Research Symposium*. The World Bank, pp. 21–42.
- Kelly, S. and Pollitt, M. (Nov. 2010). 'An Assessment of the Present and Future Opportunities for Combined Heat and Power with District Heating (CHP-DH) in the United Kingdom'. In: *Energy Policy*. Energy Efficiency Policies and Strategies with Regular Papers. 38.11, pp. 6936–6945. ISSN: 0301-4215. DOI: [10.1016/j.enpol.2010.07.010](https://doi.org/10.1016/j.enpol.2010.07.010).

- Khair, R. and Haouari, M. (Dec. 2015). 'Optimization Models for a Single-Plant District Cooling System'. In: *European Journal of Operational Research* 247.2, pp. 648–658. ISSN: 0377-2217. DOI: [10.1016/j.ejor.2015.05.083](https://doi.org/10.1016/j.ejor.2015.05.083).
- Khodabakhsh, R. and Sirouspour, S. (July 2016). 'Optimal Control of Energy Storage in a Microgrid by Minimizing Conditional Value-at-Risk'. In: *IEEE Transactions on Sustainable Energy* 7.3, pp. 1264–1273. ISSN: 1949-3029. DOI: [10.1109/TSTE.2016.2543024](https://doi.org/10.1109/TSTE.2016.2543024).
- Kleinbach, E. M., Beckman, W. A. and Klein, S. A. (Feb. 1993). 'Performance Study of One-Dimensional Models for Stratified Thermal Storage Tanks'. In: *Solar Energy* 50.2, pp. 155–166. ISSN: 0038-092X. DOI: [10.1016/0038-092X\(93\)90087-5](https://doi.org/10.1016/0038-092X(93)90087-5).
- Koltsaklis, N. E., Kopanos, G. M. and Georgiadis, M. C. (Nov. 2014). 'Design and Operational Planning of Energy Networks Based on Combined Heat and Power Units'. In: *Industrial & Engineering Chemistry Research* 53.44, pp. 16905–16923. ISSN: 0888-5885. DOI: [10.1021/ie404165c](https://doi.org/10.1021/ie404165c).
- Kopanos, G. M., Georgiadis, M. C. and Pistikopoulos, E. N. (Jan. 2013). 'Scheduling Energy Cogeneration Units under Energy Demand Uncertainty'. In: *IFAC Proceedings Volumes*. 7th IFAC Conference on Manufacturing Modelling, Management, and Control 46.9, pp. 1280–1285. ISSN: 1474-6670. DOI: [10.3182/20130619-3-RU-3018.00275](https://doi.org/10.3182/20130619-3-RU-3018.00275).
- Kotzur, L. *et al.* (Mar. 2018a). 'Impact of Different Time Series Aggregation Methods on Optimal Energy System Design'. In: *Renewable Energy* 117, pp. 474–487. ISSN: 0960-1481. DOI: [10.1016/j.renene.2017.10.017](https://doi.org/10.1016/j.renene.2017.10.017).
- (Mar. 2018b). 'Time Series Aggregation for Energy System Design: Modeling Seasonal Storage'. In: *Applied Energy* 213, pp. 123–135. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2018.01.023](https://doi.org/10.1016/j.apenergy.2018.01.023).
- Kraft, D. *et al.* (1988). *A Software Package for Sequential Quadratic Programming*. Tech. rep. DFVLR-FB-88-28. DFVLR Obersfafeuhofen, Germany. URL: <http://www.opengrey.eu/item/display/10068/147127> (visited on 23/11/2016).
- Kuznetsova, E. *et al.* (2014). 'An Integrated Framework of Agent-Based Modelling and Robust Optimization for Microgrid Energy Management'. In: *Applied Energy* 129, pp. 70–88. DOI: [10.1016/j.apenergy.2014.04.024](https://doi.org/10.1016/j.apenergy.2014.04.024).
- Land, A. H. and Doig, A. G. (1960). 'An Automatic Method of Solving Discrete Programming Problems'. In: *Econometrica* 28.3, pp. 497–520. ISSN: 0012-9682. DOI: [10.2307/1910129](https://doi.org/10.2307/1910129).
- Lee, J. H., Kim, H. and Song, Y.-h. (Jan. 2018). 'A Study on Verification of Changes in Performance of a Water-Cooled VRF System with Control Change Based on Measuring Data'. In: *Energy and Buildings* 158, pp. 712–720. ISSN: 0378-7788. DOI: [10.1016/j.enbuild.2017.10.014](https://doi.org/10.1016/j.enbuild.2017.10.014).
- Li, L. *et al.* (Aug. 2016a). 'Economic and Environmental Optimization for Distributed Energy Resource Systems Coupled with District Energy Networks'. In: *Energy* 109, pp. 947–960. ISSN: 0360-5442. DOI: [10.1016/j.energy.2016.05.026](https://doi.org/10.1016/j.energy.2016.05.026).
- Li, Z. *et al.* (Jan. 2016b). 'Combined Heat and Power Dispatch Considering Pipeline Energy Storage of District Heating Network'. In: *IEEE Transactions on Sustainable Energy* 7.1, pp. 12–22. ISSN: 1949-3029. DOI: [10.1109/TSTE.2015.2467383](https://doi.org/10.1109/TSTE.2015.2467383).
- Li, Z., Ding, R. and Floudas, C. A. (Sept. 2011). 'A Comparative Theoretical and Computational Study on Robust Counterpart Optimization: I. Robust Linear Optimization and Robust Mixed Integer Linear Optimization'. In: *Industrial & Engineering Chemistry Research* 50.18, pp. 10567–10603. ISSN: 0888-5885. DOI: [10.1021/ie200150p](https://doi.org/10.1021/ie200150p).
- Lu, D., Bao, Z. and Li, Z. (July 2016). 'Load Sampling for Scuc Based on Principal Component Analysis and Kernel Density Estimation'. In: *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pp. 1–5. DOI: [10.1109/PESGM.2016.7741946](https://doi.org/10.1109/PESGM.2016.7741946).
- Majewski, D. *et al.* (2017). 'Robust Multi-Objective Optimization for Sustainable Design of Distributed Energy Supply Systems'. In: *Computers and Chemical Engineering* 102, pp. 26–39. DOI: [10.1016/j.compchemeng.2016.11.038](https://doi.org/10.1016/j.compchemeng.2016.11.038).

- Mallipeddi, R. *et al.* (Aug. 2012). 'Efficient Constraint Handling for Optimal Reactive Power Dispatch Problems'. In: *Swarm and Evolutionary Computation* 5, pp. 28–36. ISSN: 2210-6502. DOI: [10.1016/j.swevo.2012.03.001](https://doi.org/10.1016/j.swevo.2012.03.001).
- Mancarella, P. (Feb. 2014). 'MES (Multi-Energy Systems): An Overview of Concepts and Evaluation Models'. en. In: *Energy* 65, pp. 1–17. ISSN: 03605442. DOI: [10.1016/j.energy.2013.10.041](https://doi.org/10.1016/j.energy.2013.10.041).
- Maréchal, F., Weber, C. and Favrat, D. (2008). 'Multiobjective Design and Optimization of Urban Energy Systems'. en. In: *Process Systems Engineering*. Ed. by M. C. Georgiadis, E. S. Kikkinides and E. N. Pistikopoulos. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 39–83. ISBN: 978-3-527-63129-2. URL: <http://onlinelibrary.wiley.com/doi/10.1002/9783527631292.ch2/summary> (visited on 11/11/2015).
- Marnay, C. *et al.* (Sept. 2015). 'Microgrid Evolution Roadmap'. In: *2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST)*, pp. 139–144. DOI: [10.1109/SEDST.2015.7315197](https://doi.org/10.1109/SEDST.2015.7315197).
- Marquant, J. F. *et al.* (Dec. 2017). 'A Holarchic Approach for Multi-Scale Distributed Energy System Optimisation'. In: *Applied Energy* 208, pp. 935–953. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2017.09.057](https://doi.org/10.1016/j.apenergy.2017.09.057).
- Martin, M. and Thornley, P. (2013). *The Potential for Thermal Storage to Reduce the Overall Carbon Emissions from District Heating Systems*. Tech. rep. Tyndall Centre Working Paper 157. URL: <http://www.tyndall.ac.uk/sites/default/files/twp157.pdf> (visited on 02/02/2016).
- Maurovich-Horvat, L., Rocha, P. and Siddiqui, A. S. (Jan. 2016). 'Optimal Operation of Combined Heat and Power under Uncertainty and Risk Aversion'. In: *Energy and Buildings* 110, pp. 415–425. ISSN: 0378-7788. DOI: [10.1016/j.enbuild.2015.11.009](https://doi.org/10.1016/j.enbuild.2015.11.009).
- Mavromatidis, G., Orehounig, K. and Carmeliet, J. (2018a). 'Design of Distributed Energy Systems under Uncertainty: A Two-Stage Stochastic Programming Approach'. In: *Applied Energy* 222, pp. 932–950. DOI: [10.1016/j.apenergy.2018.04.019](https://doi.org/10.1016/j.apenergy.2018.04.019).
- (2018b). 'Uncertainty and Global Sensitivity Analysis for the Optimal Design of Distributed Energy Systems'. In: *Applied Energy* 214, pp. 219–238. DOI: [10.1016/j.apenergy.2018.01.062](https://doi.org/10.1016/j.apenergy.2018.01.062).
- Mavromatidis, G., Orehounig, K. and Carmeliet, J. (Aug. 2018c). 'Comparison of Alternative Decision-Making Criteria in a Two-Stage Stochastic Program for the Design of Distributed Energy Systems under Uncertainty'. In: *Energy* 156, pp. 709–724. ISSN: 0360-5442. DOI: [10.1016/j.energy.2018.05.081](https://doi.org/10.1016/j.energy.2018.05.081).
- McKinney, W. (2010). 'Data Structures for Statistical Computing in Python'. In: *Proceedings of the 9th Python in Science Conference*. Vol. 445. Austin, TX, pp. 51–56.
- Megiddo, N. (1986). *On the Complexity of Linear Programming*. IBM Thomas J. Watson Research Division.
- Mehleri, E. D. *et al.* (Aug. 2012). 'A Mathematical Programming Approach for Optimal Design of Distributed Energy Systems at the Neighbourhood Level'. In: *Energy. Integration and Energy System Engineering, European Symposium on Computer-Aided Process Engineering 2011* 44.1, pp. 96–104. ISSN: 0360-5442. DOI: [10.1016/j.energy.2012.02.009](https://doi.org/10.1016/j.energy.2012.02.009).
- Meurant, G. (June 2014). *Algorithms and Complexity*. en. Elsevier. ISBN: 978-0-08-093391-7.
- Milan, C. *et al.* (June 2015). 'Modeling of Non-Linear CHP Efficiency Curves in Distributed Energy Systems'. In: *Applied Energy* 148, pp. 334–347. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2015.03.053](https://doi.org/10.1016/j.apenergy.2015.03.053).
- Mitchell, J. E. (2002). 'Branch-and-Cut Algorithms for Combinatorial Optimization Problems'. In: *Handbook of applied optimization*, pp. 65–77.
- Moore, C. and Mertens, S. (Aug. 2011). *The Nature of Computation*. Oxford, New York: Oxford University Press. ISBN: 978-0-19-923321-2.
- Morrison, R. (Apr. 2018). 'Energy System Modeling: Public Transparency, Scientific Reproducibility, and Open Development'. In: *Energy Strategy Reviews* 20, pp. 49–63. ISSN: 2211-467X. DOI: [10.1016/j.esr.2017.12.010](https://doi.org/10.1016/j.esr.2017.12.010).

- Morvaj, B., Evins, R. and Carmeliet, J. (Dec. 2016). 'Optimising Urban Energy Systems: Simultaneous System Sizing, Operation and District Heating Network Layout'. In: *Energy* 116, Part 1, pp. 619–636. ISSN: 0360-5442. DOI: [10.1016/j.energy.2016.09.139](https://doi.org/10.1016/j.energy.2016.09.139).
- Nahmmacher, P. *et al.* (Oct. 2016). 'Carpe Diem: A Novel Approach to Select Representative Days for Long-Term Power System Modeling'. In: *Energy* 112, pp. 430–442. ISSN: 0360-5442. DOI: [10.1016/j.energy.2016.06.081](https://doi.org/10.1016/j.energy.2016.06.081).
- Narayan, A. and Ponnambalam, K. (Feb. 2017). 'Risk-Averse Stochastic Programming Approach for Microgrid Planning under Uncertainty'. In: *Renewable Energy* 101, pp. 399–408. ISSN: 0960-1481. DOI: [10.1016/j.renene.2016.08.064](https://doi.org/10.1016/j.renene.2016.08.064).
- Neut, A. van der (2018). *Ruamel.Yaml: Ruamel.Yaml Is a YAML Parser/Emitter That Supports Roundtrip Preservation of Comments, Seq/Map Flow Style, and Map Key Order*. URL: <https://bitbucket.org/ruamel/yaml> (visited on 29/11/2018).
- Nielsen, J. (2005). *Two-Step Decision and Optimisation Model for Centralised or Decentralised Thermal Storage in DH&C Systems*. Tech. rep. 8DHC-05.02. IEA Annex VII.
- EERE – Office of Energy Efficiency and Renewable Energy (2013). *Commercial and Residential Hourly Load Profiles for All TMY3 Locations in the United States*. Tech. rep. United States Department of Energy. URL: <http://en.openei.org/doe-opendata/dataset/commercial-and-residential-hourly-load-profiles-for-all-tmy3-locations-in-the-united-states> (visited on 08/03/2017).
- Olsthoorn, D., Haghghat, F. and Mirzaei, P. A. (Oct. 2016). 'Integration of Storage and Renewable Energy into District Heating Systems: A Review of Modelling and Optimization'. In: *Solar Energy* 136, pp. 49–64. ISSN: 0038-092X. DOI: [10.1016/j.solener.2016.06.054](https://doi.org/10.1016/j.solener.2016.06.054).
- Omu, A., Choudhary, R. and Boies, A. (Oct. 2013). 'Distributed Energy Resource System Optimisation Using Mixed Integer Linear Programming'. en. In: *Energy Policy* 61, pp. 249–266. ISSN: 03014215. DOI: [10.1016/j.enpol.2013.05.009](https://doi.org/10.1016/j.enpol.2013.05.009).
- Omu, A. *et al.* (2015). 'Economic, Climate Change, and Air Quality Analysis of Distributed Energy Resource Systems'. en. In: *Procedia Computer Science* 51, pp. 2147–2156. ISSN: 18770509. DOI: [10.1016/j.procs.2015.05.487](https://doi.org/10.1016/j.procs.2015.05.487).
- OpenStreetMap contributors (2017). 'Planet Dump Retrieved from <https://planet.osm.org>'. In:
- Orehounig, K., Evins, R. and Dorer, V. (Sept. 2015). 'Integration of Decentralized Energy Systems in Neighbourhoods Using the Energy Hub Approach'. In: *Applied Energy* 154, pp. 277–289. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2015.04.114](https://doi.org/10.1016/j.apenergy.2015.04.114).
- Pazouki, S. and Haghifam, M.-R. (Sept. 2016). 'Optimal Planning and Scheduling of Energy Hub in Presence of Wind, Storage and Demand Response under Uncertainty'. In: *International Journal of Electrical Power & Energy Systems* 80, pp. 219–239. ISSN: 0142-0615. DOI: [10.1016/j.ijepes.2016.01.044](https://doi.org/10.1016/j.ijepes.2016.01.044).
- Pazouki, S., Haghifam, M.-R. and Moser, A. (Oct. 2014). 'Uncertainty Modeling in Optimal Operation of Energy Hub in Presence of Wind, Storage and Demand Response'. en. In: *International Journal of Electrical Power & Energy Systems* 61, pp. 335–345. ISSN: 01420615. DOI: [10.1016/j.ijepes.2014.03.038](https://doi.org/10.1016/j.ijepes.2014.03.038).
- Pedregosa, F. *et al.* (Oct. 2011). 'Scikit-Learn: Machine Learning in Python'. In: *Journal of Machine Learning Research* 12, pp. 2825–2830. ISSN: 1533-7928. URL: <http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html> (visited on 09/08/2018).
- Pepermans, G. *et al.* (Apr. 2005). 'Distributed Generation: Definition, Benefits and Issues'. en. In: *Energy Policy* 33.6, pp. 787–798. ISSN: 03014215. DOI: [10.1016/j.enpol.2003.10.004](https://doi.org/10.1016/j.enpol.2003.10.004).
- Pfenninger, S. (July 2017a). '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'. In: *Applied Energy* 197, pp. 1–13. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2017.03.051](https://doi.org/10.1016/j.apenergy.2017.03.051).
- (Feb. 2017b). 'Energy Scientists Must Show Their Workings'. en. In: *Nature News* 542.7642, p. 393. DOI: [10.1038/542393a](https://doi.org/10.1038/542393a).

- Pfenninger, S. and Keirstead, J. (July 2015a). 'Comparing Concentrating Solar and Nuclear Power as Baseload Providers Using the Example of South Africa'. In: *Energy* 87, pp. 303–314. ISSN: 0360-5442. DOI: [10.1016/j.energy.2015.04.077](https://doi.org/10.1016/j.energy.2015.04.077).
- (Aug. 2015b). 'Renewables, Nuclear, or Fossil Fuels? Scenarios for Great Britain's Power System Considering Costs, Emissions and Energy Security'. In: *Applied Energy* 152, pp. 83–93. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2015.04.102](https://doi.org/10.1016/j.apenergy.2015.04.102).
- Pfenninger, S. and Pickering, B. (Sept. 2018). 'Calliope: A Multi-Scale Energy Systems Modelling Framework'. en. In: *The Journal of Open Source Software* 3.29, p. 825. DOI: [10.21105/joss.00825](https://doi.org/10.21105/joss.00825).
- Pfenninger, S. and Staffell, I. (Nov. 2016). 'Long-Term Patterns of European PV Output Using 30 Years of Validated Hourly Reanalysis and Satellite Data'. In: *Energy* 114, pp. 1251–1265. ISSN: 0360-5442. DOI: [10.1016/j.energy.2016.08.060](https://doi.org/10.1016/j.energy.2016.08.060).
- Pfenninger, S. *et al.* (Feb. 2017). 'The Importance of Open Data and Software: Is Energy Research Lagging Behind?' In: *Energy Policy* 101, pp. 211–215. ISSN: 0301-4215. DOI: [10.1016/j.enpol.2016.11.046](https://doi.org/10.1016/j.enpol.2016.11.046).
- Pfenninger, S. *et al.* (Jan. 2018). 'Opening the Black Box of Energy Modelling: Strategies and Lessons Learned'. In: *Energy Strategy Reviews* 19, pp. 63–71. ISSN: 2211-467X. DOI: [10.1016/j.esr.2017.12.002](https://doi.org/10.1016/j.esr.2017.12.002).
- Pickering, B. and Choudhary, R. (Feb. 2019). 'District Energy System Optimisation under Uncertain Demand: Handling Data-Driven Stochastic Profiles'. In: *Applied Energy* 236, pp. 1138–1157. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2018.12.037](https://doi.org/10.1016/j.apenergy.2018.12.037).
- Pickering, B. and Choudhary, R. (July 2017). 'Applying Piecewise Linear Characteristic Curves in District Energy Optimisation'. In: *Proceedings of the 30th International Conference on Efficiency, Cost, Optimisation, Simulation and Environmental Impact of Energy Systems*. San Diego, USA, pp. 1080–1092.
- (Sept. 2018). 'Mitigating Risk in District-Level Energy Investment Decisions by Scenario Optimisation'. en. In: *Proceedings of BSO 2018*. Cambridge, UK, pp. 38–45. URL: <http://www.ibpsa.org/proceedings/BSO2018/1B-1.pdf>.
- Pickering, B. *et al.* (2016). 'Comparison of Metaheuristic and Linear Programming Models for the Purpose of Optimising Building Energy Supply Operation Schedule'. In: *12th REHVA World Congress*. Vol. 6. Aalborg, Denmark: Aalborg University, Department of Civil Engineering. ISBN: 87-91606-31-4. URL: [http://vbn.aau.dk/files/233775414/paper\\_529.pdf](http://vbn.aau.dk/files/233775414/paper_529.pdf) (visited on 11/08/2016).
- Plotly – Plotly Technologies Inc. (2015). *Collaborative Data Science*. Montreal, QC. URL: <https://plot.ly>.
- Powell, K. M. *et al.* (Feb. 2013). 'Optimal Chiller Loading in a District Cooling System with Thermal Energy Storage'. In: *Energy* 50, pp. 445–453. ISSN: 0360-5442. DOI: [10.1016/j.energy.2012.10.058](https://doi.org/10.1016/j.energy.2012.10.058).
- Quarton, C. J. and Samsatli, S. (Dec. 2018). 'Power-to-Gas for Injection into the Gas Grid: What Can We Learn from Real-Life Projects, Economic Assessments and Systems Modelling?' In: *Renewable and Sustainable Energy Reviews* 98, pp. 302–316. ISSN: 1364-0321. DOI: [10.1016/j.rser.2018.09.007](https://doi.org/10.1016/j.rser.2018.09.007).
- Ramsay, J. and Silverman, B. W. (2005). *Functional Data Analysis*. en. 2nd ed. Springer Series in Statistics. New York: Springer-Verlag. ISBN: 978-0-387-40080-8. URL: [//www.springer.com/gb/book/9780387400808](http://www.springer.com/gb/book/9780387400808) (visited on 28/08/2018).
- Read, D. (2005). 'Some Observations on Resilience and Robustness in Human Systems'. In: *Cybernetics and Systems: An International Journal* 36.8, pp. 773–802.
- Reddy, S. (2017). 'Optimization of Renewable Energy Resources in Hybrid Energy Systems'. In: *Journal of Green Engineering* 7.1-2, pp. 43–60. DOI: [10.13052/jge1904-4720.7123](https://doi.org/10.13052/jge1904-4720.7123).
- Refaeilzadeh, P., Tang, L. and Liu, H. (2009). 'Cross-Validation'. en. In: *Encyclopedia of Database Systems*. Springer, Boston, MA, pp. 532–538. DOI: [10.1007/978-0-387-39940-9\\_565](https://doi.org/10.1007/978-0-387-39940-9_565). URL: [https://link.springer.com/referenceworkentry/10.1007/978-0-387-39940-9\\_565](https://link.springer.com/referenceworkentry/10.1007/978-0-387-39940-9_565) (visited on 09/08/2018).
- Reinhart, C. F. and Cerezo Davila, C. (Feb. 2016). 'Urban Building Energy Modeling – A Review of a Nascent Field'. In: *Building and Environment* 97, pp. 196–202. ISSN: 0360-1323. DOI: [10.1016/j.buildenv.2015.12.001](https://doi.org/10.1016/j.buildenv.2015.12.001).

- Ren, H. *et al.* (Dec. 2010). 'Multi-Objective Optimization for the Operation of Distributed Energy Systems Considering Economic and Environmental Aspects'. en. In: *Applied Energy* 87.12, pp. 3642–3651. ISSN: 03062619. DOI: [10.1016/j.apenergy.2010.06.013](https://doi.org/10.1016/j.apenergy.2010.06.013).
- Rezaie, B. and Rosen, M. A. (May 2012). 'District Heating and Cooling: Review of Technology and Potential Enhancements'. In: *Applied Energy*. (1) Green Energy; (2) Special Section from Papers Presented at the 2nd International Energy 2030 Conf 93, pp. 2–10. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2011.04.020](https://doi.org/10.1016/j.apenergy.2011.04.020).
- Rezvan, A. T., Gharneh, N. S. and Gharehpetian, G. B. (Feb. 2013). 'Optimization of Distributed Generation Capacities in Buildings under Uncertainty in Load Demand'. In: *Energy and Buildings* 57, pp. 58–64. ISSN: 0378-7788. DOI: [10.1016/j.enbuild.2012.10.031](https://doi.org/10.1016/j.enbuild.2012.10.031).
- Rockafellar, R. T. and Uryasev, S. (July 2002). 'Conditional Value-at-Risk for General Loss Distributions'. In: *Journal of Banking & Finance* 26.7, pp. 1443–1471. ISSN: 0378-4266. DOI: [10.1016/S0378-4266\(02\)00271-6](https://doi.org/10.1016/S0378-4266(02)00271-6).
- Römisch, W. (2009). 'Scenario Reduction Techniques in Stochastic Programming'. In: *Stochastic Algorithms: Foundations and Applications*. Ed. by O. Watanabe and T. Zeugmann. Vol. 5792. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–14. ISBN: 978-3-642-04943-9. DOI: [10.1007/978-3-642-04944-6\\_1](https://doi.org/10.1007/978-3-642-04944-6_1). URL: [http://link.springer.com/10.1007/978-3-642-04944-6\\_1](http://link.springer.com/10.1007/978-3-642-04944-6_1) (visited on 17/01/2018).
- Rosen, M. A. (2001). 'The Exergy of Stratified Thermal Energy Storages'. In: *Solar Energy* 71.3, pp. 173–185. ISSN: 0038-092X. DOI: [10.1016/S0038-092X\(01\)00036-6](https://doi.org/10.1016/S0038-092X(01)00036-6).
- Rotman, B. (Aug. 2003). 'Will the Digital Computer Transform Classical Mathematics?' en. In: *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 361.1809, pp. 1675–1690. ISSN: 1364-503X, 1471-2962. DOI: [10.1098/rsta.2003.1230](https://doi.org/10.1098/rsta.2003.1230).
- Sameti, M. and Haghghat, F. (Apr. 2017). 'Optimization Approaches in District Heating and Cooling Thermal Network'. In: *Energy and Buildings* 140, pp. 121–130. ISSN: 0378-7788. DOI: [10.1016/j.enbuild.2017.01.062](https://doi.org/10.1016/j.enbuild.2017.01.062).
- Sharafi, M. and ElMekkawy, T. Y. (Aug. 2014). 'Multi-Objective Optimal Design of Hybrid Renewable Energy Systems Using PSO-Simulation Based Approach'. en. In: *Renewable Energy* 68, pp. 67–79. ISSN: 09601481. DOI: [10.1016/j.renene.2014.01.011](https://doi.org/10.1016/j.renene.2014.01.011).
- (Dec. 2015). 'Stochastic Optimization of Hybrid Renewable Energy Systems Using Sampling Average Method'. In: *Renewable and Sustainable Energy Reviews* 52, pp. 1668–1679. ISSN: 1364-0321. DOI: [10.1016/j.rser.2015.08.010](https://doi.org/10.1016/j.rser.2015.08.010).
- Siriruk, P. (2012). 'Fitting Piecewise Linear Functions Using Particle Swarm Optimization'. In: *Suranaree J. Sci. Technol* 19.4, pp. 259–264. URL: <http://ird.sut.ac.th/e-journal/Journal/pdf/1301004.pdf> (visited on 23/11/2016).
- Smith, J. C. and Taskin, Z. C. (2008). 'A Tutorial Guide to Mixed-Integer Programming Models and Solution Techniques'. In: *Optimization in Medicine and Biology*, pp. 521–548. URL: [http://www.ie.boun.edu.tr/~taskin/pdf/IP\\_tutorial.pdf](http://www.ie.boun.edu.tr/~taskin/pdf/IP_tutorial.pdf) (visited on 17/11/2015).
- Söderman, J. (Nov. 2007). 'Optimisation of Structure and Operation of District Cooling Networks in Urban Regions'. In: *Applied Thermal Engineering*. Selected Papers from the 9th Conference on Process Integration, Modelling and Optimisation for Energy Saving and Pollution Reduction – PRES2006 27.16, pp. 2665–2676. ISSN: 1359-4311. DOI: [10.1016/j.applthermaleng.2007.05.004](https://doi.org/10.1016/j.applthermaleng.2007.05.004).
- Söderman, J. and Pettersson, F. (Sept. 2006). 'Structural and Operational Optimisation of Distributed Energy Systems'. en. In: *Applied Thermal Engineering* 26.13, pp. 1400–1408. ISSN: 13594311. DOI: [10.1016/j.applthermaleng.2005.05.034](https://doi.org/10.1016/j.applthermaleng.2005.05.034).
- Srikhirin, P., Aphornratana, S. and Chungpaibulpatana, S. (Dec. 2001). 'A Review of Absorption Refrigeration Technologies'. In: *Renewable and Sustainable Energy Reviews* 5.4, pp. 343–372. ISSN: 1364-0321. DOI: [10.1016/S1364-0321\(01\)00003-X](https://doi.org/10.1016/S1364-0321(01)00003-X).

- Staffell, I. and Pfenninger, S. (Nov. 2016). 'Using Bias-Corrected Reanalysis to Simulate Current and Future Wind Power Output'. In: *Energy* 114, pp. 1224–1239. ISSN: 0360-5442. DOI: [10.1016/j.energy.2016.08.068](https://doi.org/10.1016/j.energy.2016.08.068).
- Steen, D. *et al.* (Jan. 2015). 'Modeling of Thermal Storage Systems in MILP Distributed Energy Resource Models'. en. In: *Applied Energy* 137, pp. 782–792. ISSN: 03062619. DOI: [10.1016/j.apenergy.2014.07.036](https://doi.org/10.1016/j.apenergy.2014.07.036).
- Sun, Y. (Aug. 2014). 'Closing the Building Energy Performance Gap by Improving Our Predictions'. en\_US. Doctor of Philosophy. Georgia Institute of Technology. URL: <https://smartech.gatech.edu/handle/1853/52285> (visited on 17/01/2018).
- Swan, L. G. and Ugursal, V. I. (Oct. 2009). 'Modeling of End-Use Energy Consumption in the Residential Sector: A Review of Modeling Techniques'. In: *Renewable and Sustainable Energy Reviews* 13.8, pp. 1819–1835. ISSN: 1364-0321. DOI: [10.1016/j.rser.2008.09.033](https://doi.org/10.1016/j.rser.2008.09.033).
- Takahama, T. and Sakai, S. (July 2010). 'Constrained Optimization by the #x03B5; Constrained Differential Evolution with an Archive and Gradient-Based Mutation'. In: *IEEE Congress on Evolutionary Computation*, pp. 1–9. DOI: [10.1109/CEC.2010.5586484](https://doi.org/10.1109/CEC.2010.5586484).
- Taljan, G. *et al.* (May 2012). 'Optimal Sizing of Biomass-Fired Organic Rankine Cycle CHP System with Heat Storage'. In: *Renewable Energy* 41, pp. 29–38. ISSN: 0960-1481. DOI: [10.1016/j.renene.2011.09.034](https://doi.org/10.1016/j.renene.2011.09.034).
- Tanaka, I., Yuge, H. and Ohmori, H. (2017). 'Formulation and Evaluation of Long-Term Allocation Problem for Renewable Distributed Generations'. In: *IET Renewable Power Generation* 11.12, pp. 1584–1596. ISSN: 1752-1416. DOI: [10.1049/iet-rpg.2017.0068](https://doi.org/10.1049/iet-rpg.2017.0068).
- Taylor, A. (2017). 'Why the Social Cost of Carbon Is a Red Herring Comments'. eng. In: *Tulane Environmental Law Journal* 31, pp. 345–372. URL: <https://heinonline.org/HOL/P?h=hein:journals/tulev31%5C&i=371> (visited on 20/09/2018).
- The Carbon Trust (July 2011). *Closing the Gap: Lessons Learned on Realising the Potential of Low Carbon Building Design*. Tech. rep. The Carbon Trust.
- BEIS – The Department for Business, Energy and Industrial Strategy (Jan. 2018b). *Updated Short Term Traded Carbon Values*. Tech. rep. URL: <https://www.gov.uk/government/publications/updated-short-term-traded-carbon-values-used-for-modelling-purposes-2017> (visited on 20/09/2018).
- The Mathworks, Inc. (2015). *MATLAB and Parallel Computing Toolbox*. Natick, MA, US.
- YAML – *The Official YAML Web Site* (2009). URL: <http://yaml.org/> (visited on 29/11/2018).
- The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE) (2003). *Computer Aided Simulation for Cogeneration Assessment & Design III*. Tokyo: Maruzen publishing.
- Turlach, B. A. (1993). 'Bandwidth Selection in Kernel Density Estimation: A Review'. In: *CORE and Institut de Statistique*. Vol. 19, pp. 1–33.
- United Nations Environment Programme (2014). *District Energy in Cities*. Tech. rep. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.679.1556%5C&rep=rep1%5C&type=pdf> (visited on 05/08/2016).
- Vahid-Pakdel, M. J. *et al.* (Aug. 2017). 'Stochastic Optimization of Energy Hub Operation with Consideration of Thermal Energy Market and Demand Response'. In: *Energy Conversion and Management* 145, pp. 117–128. ISSN: 0196-8904. DOI: [10.1016/j.enconman.2017.04.074](https://doi.org/10.1016/j.enconman.2017.04.074).
- Verrilli, F. *et al.* (Apr. 2017). 'Model Predictive Control-Based Optimal Operations of District Heating System With Thermal Energy Storage and Flexible Loads'. In: *IEEE Transactions on Automation Science and Engineering* 14.2, pp. 547–557. ISSN: 1545-5955. DOI: [10.1109/TASE.2016.2618948](https://doi.org/10.1109/TASE.2016.2618948).
- Voll, P. *et al.* (Feb. 2013). 'Automated Superstructure-Based Synthesis and Optimization of Distributed Energy Supply Systems'. In: *Energy* 50, pp. 374–388. ISSN: 0360-5442. DOI: [10.1016/j.energy.2012.10.045](https://doi.org/10.1016/j.energy.2012.10.045).
- Voll, P. *et al.* (Mar. 2015). 'The Optimum Is Not Enough: A near-Optimal Solution Paradigm for Energy Systems Synthesis'. In: *Energy* 82, pp. 446–456. ISSN: 0360-5442. DOI: [10.1016/j.energy.2015.01.055](https://doi.org/10.1016/j.energy.2015.01.055).

- Wakui, T., Kinoshita, T. and Yokoyama, R. (Apr. 2014). 'A Mixed-Integer Linear Programming Approach for Cogeneration-Based Residential Energy Supply Networks with Power and Heat Interchanges'. In: *Energy* 68, pp. 29–46. ISSN: 0360-5442. DOI: [10.1016/j.energy.2014.01.110](https://doi.org/10.1016/j.energy.2014.01.110).
- Wakui, T. and Yokoyama, R. (Jan. 2014). 'Optimal Structural Design of Residential Cogeneration Systems in Consideration of Their Operating Restrictions'. In: *Energy* 64, pp. 719–733. ISSN: 0360-5442. DOI: [10.1016/j.energy.2013.10.002](https://doi.org/10.1016/j.energy.2013.10.002).
- Wang, C. *et al.* (Jan. 2017). 'Multi-Objective Optimization and Parametric Analysis of Energy System Designs for the Albano University Campus in Stockholm'. In: *Procedia Engineering*. International High-Performance Built Environment Conference – A Sustainable Built Environment Conference 2016 Series (SBE16), iHBE 2016 180, pp. 621–630. ISSN: 1877-7058. DOI: [10.1016/j.proeng.2017.04.221](https://doi.org/10.1016/j.proeng.2017.04.221).
- Wang, H. *et al.* (Dec. 2015). 'Modelling and Optimization of CHP Based District Heating System with Renewable Energy Production and Energy Storage'. In: *Applied Energy* 159, pp. 401–421. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2015.09.020](https://doi.org/10.1016/j.apenergy.2015.09.020).
- Ward, R., Choudhary, R. and Aston, J. (Sept. 2018). 'A Stochastic Data-Centric Model for Quantification of End-Use Energy Demand in Buildings'. In: *Proceedings of BSO 2018*. Cambridge, UK, pp. 307–314. URL: <http://www.ibpsa.org/proceedings/BSO2018/3C-2.pdf> (visited on 23/10/2018).
- Ward, R. *et al.* (Sept. 2016a). 'Data Driven Bottom-up Approach for Modelling Internal Loads in Building Energy Simulation Using Functional Principal Components'. In: *Proceedings of the 3rd IBPSA-England Conference BSO 2016*. Great North Museum, Newcastle, pp. 1115–1122.
- Ward, R. *et al.* (July 2016b). 'Exploring the Impact of Different Parameterisations of Occupant-Related Internal Loads in Building Energy Simulation'. In: *Energy and Buildings* 123, pp. 92–105. ISSN: 0378-7788. DOI: [10.1016/j.enbuild.2016.04.050](https://doi.org/10.1016/j.enbuild.2016.04.050).
- Weber, C. and Shah, N. (2011). 'Optimisation Based Design of a District Energy System for an Eco-Town in the United Kingdom'. In: *Energy* 36.2, pp. 1292–1308. DOI: [10.1016/j.energy.2010.11.014](https://doi.org/10.1016/j.energy.2010.11.014).
- Wei, F. *et al.* (2016). 'Optimal Unit Sizing for Small-Scale Integrated Energy Systems Using Multi-Objective Interval Optimization and Evidential Reasoning Approach'. In: *Energy* 111, pp. 933–946. DOI: [10.1016/j.energy.2016.05.046](https://doi.org/10.1016/j.energy.2016.05.046).
- Woods, D. D. (Sept. 2015). 'Four Concepts for Resilience and the Implications for the Future of Resilience Engineering'. In: *Reliability Engineering & System Safety*. Special Issue on Resilience Engineering 141, pp. 5–9. ISSN: 0951-8320. DOI: [10.1016/j.res.2015.03.018](https://doi.org/10.1016/j.res.2015.03.018).
- Wouters, C. *et al.* (Sept. 2014). 'Mixed-Integer Optimisation Based Approach for Design and Operation of Distributed Energy Systems'. In: *Power Engineering Conference (AUPEC), 2014 Australasian Universities*, pp. 1–6. DOI: [10.1109/AUPEC.2014.6966501](https://doi.org/10.1109/AUPEC.2014.6966501).
- Wouters, C., Fraga, E. S. and James, A. M. (June 2015). 'An Energy Integrated, Multi-Microgrid, MILP (Mixed-Integer Linear Programming) Approach for Residential Distributed Energy System Planning – A South Australian Case-Study'. In: *Energy* 85, pp. 30–44. ISSN: 0360-5442. DOI: [10.1016/j.energy.2015.03.051](https://doi.org/10.1016/j.energy.2015.03.051).
- Wu, J.-y., Wang, J.-l. and Li, S. (Dec. 2012). 'Multi-Objective Optimal Operation Strategy Study of Micro-CCHP System'. In: *Energy*. 6th Dubrovnik Conference on Sustainable Development of Energy Water and Environmental Systems, SDEWES 2011 48.1, pp. 472–483. ISSN: 0360-5442. DOI: [10.1016/j.energy.2012.10.013](https://doi.org/10.1016/j.energy.2012.10.013).
- Xia, J. *et al.* (2002). 'Testing Methodology For VRF Systems'. en. In: *International Refrigeration and Air Conditioning Conference*. Vol. Paper 542. Purdue University, p. 9.
- Yang, Y., Zhang, S. and Xiao, Y. (Jan. 2017). 'Optimal Design of Distributed Energy Resource Systems Based on Two-Stage Stochastic Programming'. In: *Applied Thermal Engineering* 110, pp. 1358–1370. ISSN: 1359-4311. DOI: [10.1016/j.applthermaleng.2016.09.049](https://doi.org/10.1016/j.applthermaleng.2016.09.049).

- Yazdanie, M., Densing, M. and Wokaun, A. (Sept. 2016). 'The Role of Decentralized Generation and Storage Technologies in Future Energy Systems Planning for a Rural Agglomeration in Switzerland'. In: *Energy Policy* 96, pp. 432–445. ISSN: 0301-4215. DOI: [10.1016/j.enpol.2016.06.010](https://doi.org/10.1016/j.enpol.2016.06.010).
- Yokoyama, R. and Ito, K. (Mar. 2002). 'Optimal Design of Energy Supply Systems Based on Relative Robustness Criterion'. In: *Energy Conversion and Management* 43.4, pp. 499–514. ISSN: 0196-8904. DOI: [10.1016/S0196-8904\(01\)00027-9](https://doi.org/10.1016/S0196-8904(01)00027-9).
- Yokoyama, R., Ohkura, M. and Wakui, T. (Sept. 2014). 'Robust Optimal Operation of a Gas Turbine Cogeneration Plant Under Uncertain Energy Demands'. In: *Journal of Engineering for Gas Turbines and Power* 137.2. ISSN: 0742-4795. DOI: [10.1115/1.4028211](https://doi.org/10.1115/1.4028211).
- Yokoyama, R. *et al.* (June 2016). 'Evaluation of Performance Robustness of a Gas Turbine Cogeneration Plant Based on a Mixed-Integer Linear Model'. In: V003T20A003. DOI: [10.1115/GT2016-56217](https://doi.org/10.1115/GT2016-56217).
- Zangeneh, A., Jadid, S. and Rahimi-Kian, A. (Dec. 2009). 'Promotion Strategy of Clean Technologies in Distributed Generation Expansion Planning'. In: *Renewable Energy* 34.12, pp. 2765–2773. ISSN: 0960-1481. DOI: [10.1016/j.renene.2009.06.018](https://doi.org/10.1016/j.renene.2009.06.018).
- Zhang, D. *et al.* (June 2015a). 'Optimal Design of CHP-Based Microgrids: Multiobjective Optimisation and Life Cycle Assessment'. In: *Energy* 85, pp. 181–193. ISSN: 0360-5442. DOI: [10.1016/j.energy.2015.03.036](https://doi.org/10.1016/j.energy.2015.03.036).
- Zhang, O., Yu, S. and Liu, P. (Apr. 2015b). 'Development Mode for Renewable Energy Power in China: Electricity Pool and Distributed Generation Units'. In: *Renewable and Sustainable Energy Reviews* 44, pp. 657–668. ISSN: 13640321. DOI: [10.1016/j.rser.2015.01.020](https://doi.org/10.1016/j.rser.2015.01.020).
- Zhao, H.-x. and Magoulès, F. (Aug. 2012). 'A Review on the Prediction of Building Energy Consumption'. In: *Renewable and Sustainable Energy Reviews* 16.6, pp. 3586–3592. ISSN: 1364-0321. DOI: [10.1016/j.rser.2012.02.049](https://doi.org/10.1016/j.rser.2012.02.049).
- Zheng, Y. *et al.* (Aug. 2018). 'Optimization under Uncertainty of a Biomass-Integrated Renewable Energy Microgrid with Energy Storage'. In: *Renewable Energy* 123, pp. 204–217. ISSN: 0960-1481. DOI: [10.1016/j.renene.2018.01.120](https://doi.org/10.1016/j.renene.2018.01.120).
- Zhou, Y. *et al.* (Apr. 2018). 'A Robust Optimization Approach for Integrated Community Energy System in Energy and Ancillary Service Markets'. In: *Energy* 148, pp. 1–15. ISSN: 0360-5442. DOI: [10.1016/j.energy.2018.01.078](https://doi.org/10.1016/j.energy.2018.01.078).
- Zhou, Z. *et al.* (Mar. 2013). 'A Two-Stage Stochastic Programming Model for the Optimal Design of Distributed Energy Systems'. In: *Applied Energy* 103, pp. 135–144. ISSN: 0306-2619. DOI: [10.1016/j.apenergy.2012.09.019](https://doi.org/10.1016/j.apenergy.2012.09.019).
- Zmeureanu, R. (Sept. 2002). 'Prediction of the Cop of Existing Rooftop Units Using Artificial Neural Networks and Minimum Number of Sensors'. In: *Energy* 27.9, pp. 889–904. ISSN: 0360-5442. DOI: [10.1016/S0360-5442\(02\)00027-0](https://doi.org/10.1016/S0360-5442(02)00027-0).
- Zugno, M., Morales, J. and Madsen, H. (2016). 'Commitment and Dispatch of Heat and Power Units via Affinely Adjustable Robust Optimization'. In: *Computers and Operations Research* 75, pp. 191–201. DOI: [10.1016/j.cor.2016.06.002](https://doi.org/10.1016/j.cor.2016.06.002).

# Appendix A

## State-of-the-art district energy system MILP optimisation models

Table A.1 Model reference table.

Model number	Model name	Citation
1	-	(Garcia and Weisser, 2006)
2	-	(Söderman and Pettersson, 2006)
3	-	(Söderman, 2007)
4	ICS-EM	(Cai <i>et al.</i> , 2009)
5	-	(Hawkes and Leach, 2009)
6	SynCity	(Keirstead <i>et al.</i> , 2010)
7	-	(Ren <i>et al.</i> , 2010)
8	DESDOP	(Weber and Shah, 2011)
9	-	(Christidis <i>et al.</i> , 2012)
10	-	(Mehleri <i>et al.</i> , 2012)
11	-	(Buoro <i>et al.</i> , 2013)
12	-	(Hoke <i>et al.</i> , 2013)
13	DENO	(Omu <i>et al.</i> , 2013)
14	-	(Voll <i>et al.</i> , 2013)
15	-	(Bischi <i>et al.</i> , 2014)
16	-	(Buoro <i>et al.</i> , 2014)
17	-	(Capuder and Mancarella, 2014)
18	-	(Fazlollahi <i>et al.</i> , 2014)
19	-	(Haikarainen <i>et al.</i> , 2014)
20	RESCOM	(Jennings <i>et al.</i> , 2014)
21	-	(Koltsaklis <i>et al.</i> , 2014)
22	-	(Pazouki <i>et al.</i> , 2014)
23	-	(Wakui <i>et al.</i> , 2014)
24	-	(Wouters <i>et al.</i> , 2014)
25	-	(Yokoyama <i>et al.</i> , 2014)
26	XEMS13	(Carpaneto <i>et al.</i> , 2015)

Model number	Model name	Citation
27	-	(Khir and Haouari, 2015)
28	DENO	(Omu <i>et al.</i> , 2015)
29	-	(Orehounig <i>et al.</i> , 2015)
30	eSynthesis	(Voll <i>et al.</i> , 2015)
31	DESDOP	(Voll <i>et al.</i> , 2015)
32	RESCOM	(Voll <i>et al.</i> , 2015)
33	-	(Wang <i>et al.</i> , 2015)
34	-	(Wouters <i>et al.</i> , 2015)
35	-	(Zhang <i>et al.</i> , 2015a)
36	-	(Akbari <i>et al.</i> , 2016)
37	-	(Ameri and Besharati, 2016)
38	-	(Ceseña <i>et al.</i> , 2016)
39	-	(Chen <i>et al.</i> , 2016)
40	-	(Farzaneh <i>et al.</i> , 2016)
41	DeST	(Li <i>et al.</i> , 2016a)
42	-	(Morvaj <i>et al.</i> , 2016)
43	-	(Pazouki and Haghifam, 2016)
44	-	(Yazdanie <i>et al.</i> , 2016)
45	-	(Yokoyama <i>et al.</i> , 2016)
46	-	(Al Rafea <i>et al.</i> , 2017)
47	ETEM-SG	(Babonneau <i>et al.</i> , 2017)
48	-	(Good and Mancarella, 2017)
49	-	(Majewski <i>et al.</i> , 2017)
50	-	(Reddy, 2017)
51	DGOPT	(Tanaka <i>et al.</i> , 2017)
52	-	(Vahid-Pakdel <i>et al.</i> , 2017)
53	-	(Bucciarelli <i>et al.</i> , 2018)
54	-	(Cesena and Mancarella, 2018)
55	-	(Gabrielli <i>et al.</i> , 2018)
56	-	(Kotzur <i>et al.</i> , 2018a; Kotzur <i>et al.</i> , 2018b)
57	-	(Mavromatidis <i>et al.</i> , 2018c)
58	-	(Mavromatidis <i>et al.</i> , 2018a; Mavromatidis <i>et al.</i> , 2018b)
59	-	(Zheng <i>et al.</i> , 2018)
60	-	(Zhou <i>et al.</i> , 2018)

Table A.2 Glossary of columnwise abbreviations used in Table A.3.

<b>Resolution</b>	<b>Column</b>	<b>Format</b>	<b>Explanation of abbreviations</b>
	<i>Temporal</i>	#A-#B/#C, where A = resolution, B = time horizon, C = timeseries aggregation (if applicable)	H = hour(s), min = minute(s), D = days(s), Y = year(s), P = period(s) (usually changeable number of hours), PL = peak load, TD = typical day(s), TW = typical week(s), TS = typical season(s)
	<i>Spatial</i>	#A-#B-C, where A = number of nodes, B = number of buildings represented by nodes (if applicable), C = district type	N = node, B = building, D = geographically located district, A = abstract district, +N = district network is described
<b>Optimisation</b>	<i>Mode</i>	-	O = operation, P = planning
	<i>Method</i>	A (B), where A is either MILP or LP and B is the software used (if known)	-
	<i>Objective</i>	A (B), where A is the primary objective and B are auxiliary objectives which are scaled to A with a weighting factor (e.g. cost of carbon)	£ = cost, CO <sub>2</sub> = carbon emissions, $\eta$ = energy/efficiency
<b>Model elements</b>	<i>Technologies</i>	If A-B, A = fuel type, B = technology	Many technologies link to the main thesis acronym list. NG = natural gas, O = oil, BM = biomass, W = wood, B = boiler, G = generator, DR = demand response
	<i>Energy types</i>	-	E = electricity, H = heat, C = cooling
	<i>N(on)L(inear) repr(resentation)</i>	A: B, where A = non-typical method for representing nonlinear characteristics and B = aspects of model to which non-typical linearisation is applied	P = piecewise, y-SLE = y-intercept <a href="#">single line efficiency</a> , ML = non-zero minimum load rate, OR = operating region, VD = volume discretisation, HF = heat flow, PF = power flow
	<i>Uncertainty</i>	If A: B, A = optimisation method, B = aspects of model for which uncertainty is described	S = independent scenarios, Se = sensitivity analysis, SO = <a href="#">scenario optimisation</a> , RO = <a href="#">robust optimisation</a> , CC = chance constrained programming, I = interval programming

Table A.3 Details of models studied in reviewed literature. Refer to Table A.1 for reference(s) associated with a model number and to Table A.2 for a glossary of abbreviations used in this table.

Model	Resolution		Optimisation			Model elements				Additional elements
	Temporal	Spatial	Mode	Method	Obj	Technologies	Energy types	NL repr.	Uncertainty analysis	
1	1H-1Y	D	P	MILP (CPLEX)	£	O/FC-G, WT, StoreH2	E	-	-	Compares LP and heuristic method
2	12H-1Y/4TD	12N-D+N	P	MILP (CPLEX)	£	NG-CHP, AHP, StoreT	E, H	-	-	-
3	12H-1Y/4TD	20/53B-D+N	P	MILP (CPLEX)	£	StoreT, ECh	C	£: y-SLE	-	Same district analysed for 2006 (20 nodes) and 2020 (53 nodes)
4	5Y-15Y	3N-AD+N	P	MILP	£	NGB, PV, H, WT	E	-	I & CC	Includes transport energy demand.
5	1H-1Y/12TD	3B-D+N	P	LP	£	NG-CHP, WT, PV, NGB, StoreE, StoreT	E, H	-	Se	-
6	1H-1Y	MB-D+N	P	MILP	£	NG-CHP, NGB	E, H	-	-	Endogenous Demand
7	1H-3D	D	O	MILP	£, (CO <sub>2</sub> )	PV, FC, CHP, HRAR, NGB	E, H, C	-	-	-
8	6P-1Y/3TD	35B-D+N	P	MILP (CPLEX)	£	NG-CHP, NGB, StoreT, ST, GSHP, WT	E, H, WH	-	Se	Heat pipe routing algorithm.
9	4H-0.5Y	D	P	MILP (CPLEX via GAMS)	£, (CO <sub>2</sub> )	NG-CHP, B-CHP, NGB, StoreT	E, H	OR: CHP	-	Only the StoreT is a planning decision variable
10	6P-1Y/3TD	10B-AD+N	P	MILP (CPLEX)	£	NG/FC-CHP, PV, NGB	E, H	-	-	Multiple NG-CHP types, providing range of HTPs
11	1H-1Y/12RW	9N-D+N	P	MILP	£, CO <sub>2</sub>	NG-CHP, NGB, ST, StoreT	E, H	-	S	-
12	1H-3D	AD+N	O	MILP (MATLAB)	£	O-G, PV, StoreE, DR	E	-	-	-
13	1H-1Y/4TD	6B-D+N	P	MILP (CPLEX)	£	NG-CHP, WT, NG/BM-B, PV, ST, AHP, GSHP	E, H	-	-	Particulate emission mapping.
14	1M+2PL-1Y	6B-D+N	O	MILP (CPLEX via GAMS)	£	NGB, AR, ECh, CHP	E, H, C	P: CHP,	-	Heuristic top-layer for planning.

Table A.3 continued from previous page

Model	Resolution		Optimisation			Model elements				Additional elements
	Temporal	Spatial	Mode	Method	Obj	Technologies	Energy types	NL repr.	Uncertainty analysis	
15	1H-1D	1N-D	O	MILP	£	HP, StoreT, NGB, ICE, GT	E, H	P: GT, ICE, NGB, HP	-	-
16	1H-1Y	9N-D+N	P	MILP	£	NG-CHP, NGB, AR, ECh, StoreT	E, H, C	-	-	Seasonal storage
17	1H-1D / H-1Y	1N-1000B-D	O	MILP	£	NG-CHP, NGB, AHP, StoreT	E, H	P: CHP	Se	Technology sizing by sensitivity analysis
18	1H-20Y/7TD	1N-500000B-D	O	MILP	£, CO <sub>2</sub> , $\eta$	W-HP, StoreT, NGB, NG-CHP	E, H	P: CHP; VD: StoreT	-	multi-objective metaheuristic top-layer for planning.
19	12P-1Y	26N-D+N	P	MILP (CPLEX via MATLAB)	£	NG/BM/O-B, AHP, StoreT	E(?), H	-	-	Detailed pipeline design. Time periods are day and night every two months.
20	6P-8Y/4TS	19N-D+N	P	MILP	£	NG-CHP, NGB, StoreE, GSHP, EH, ST, DR	E, H	-	S	Building retrofit options.
21	1H-1Y/4TD	4B-D	P	MILP	£	NG-CHP, FC-CHP, StoreT, NGB	E, H	-	S	Considers 8 different CHP types (engine, turbine, fuel cell, etc.). Waste heat from a refinery.
22	1H-1D	1N-AD	O	MILP	£	WT, NG-CHP, NGB, StoreT, StoreE, DR	E, H	-	SO	ex-post OOS tests
23	1H-1Y/12TD	20N-1B-D+N	O	MILP	£	NG-CHP, NGB, StoreT, FC-CHP, E-B	E, H	HF	-	Linearised temperature tracking in pipeline
24	1H-1Y/4TD	5B-AD	P	MILP (CPLEX)	£	NG/FC-CHP, NGB, StoreT, StoreE, PV, WT	E, H	-	S	Compares two FIT systems.
25	1H-1D	1N-D	O	MILP	£	WH-B, NGB, GT, ECh, AR	E, H, C	y-SLE	minmax regret	-
26	1H-1Y/3TW	1N-D	O	MILP (CPLEX via Matlab)	£	NG-CHP, NGB, StoreT, ST	H	P: CHP	-	-
27	variable (4-8 timesteps)	variable (15-60N-AD+N)	P	MILP (CPLEX)	£	ECh, StoreT	C	HF	S	Linearised temperature/pressure tracking in the system

Table A.3 continued from previous page

Model	Resolution		Optimisation			Model elements				Additional elements
	Temporal	Spatial	Mode	Method	Obj	Technologies	Energy types	NL repr.	Uncertainty analysis	
28	1H-19Y/12TD	5N-75B-D+N	P	MILP	£	B-CHP, NG-CHP, PV, AHP, NGB, BM-B, GSHP, AR, Ech	E, H, C	-	S	
29	1H-1Y	29B-D+N	O	MILP	CO <sub>2</sub>	B-CHP, O-B, PV, GSHP, StoreT, H, StoreE, ST	E, H	-	-	Planning by scenarios. Includes lifetime technology emissions
30	1M+2PL-1Y	6N-21B-D+N	P	MILP	£	NGB, Ech, AR, NG-CHP	E, H, C	P: CHP, NGB, AR, Ech	-	Collects near-optimal solutions.
31	1H-1Y	39N-D+N	P	MILP	£	NG-CHP, NGB, PV, WT, HP	E, H	-	-	-
32	6H-7Y/3TS	19N-92,170B-D+N	P	MILP	£	NG-CHP, GSHP, DR	E, H	-	-	-
33	1H-M	D	O	LP	£	NG/B-CHP, NG/O-B, ST, StoreT	E, H	-	-	seasonal temperature variation calculated ex-ante
34	1H-1Y/3TD	5B-AD+N	P	MILP (CPLEX via GAMS)	£	NG-CHP, WT, PV, StoreE, AR, Ech, StoreT	E, H, C	-	S, Se	Impact of solar energy uncertainty calculated ex-post.
35	6T-1Y/3TD	5B-AD+N	P	MILP (CPLEX via GAMS)	£, CO <sub>2</sub> , SO <sub>2</sub>	NG/FC-CHP, NGB, StoreT	E, H, WH	-	S	Includes embodied technology CO <sub>2</sub> costs, based on LCA
36	6T-1Y/3TD	5B-AD+N	P	MILP (CPLEX via GAMS)	£	PV, NGB, ST, Ech, AR, StoreT	E, H, C	-	RO: demand	Unmet demand calculated ex-post
37	1H-1Y/2TD	7N-137B-D+N	P	MILP (CPLEX via AIMMS)	£	CCHP, NGB, Ech, PV	E, H, C	-	S	Array of CCHP & NGB technology choices
38	30min-1Y/3TD	1N-1000B-D	P	MILP	£	CHP, AHP, StoreT	E, H	-	~RO	-

Table A.3 continued from previous page

Model	Resolution		Optimisation			Model elements				Additional elements
	Temporal	Spatial	Mode	Method	Obj	Technologies	Energy types	NL repr.	Uncertainty analysis	
39	5Y-15Y	1N-D	P	MILP	£, (CO <sub>2</sub> , air quality)	PV, ST, WT, Biomass, CCHP	E, H, C	-	CC	-
40	5Y-20Y	C	P	MILP (GAMS)	£	WT, PV, Hydro-/Geothermal-/waste-/fossil-power plant, GSHP, ST	E	-	Se, S	Including transport demand & e-constrained CO <sub>2</sub> emissions
41	2H-1Y/3TD	4B-AD+N	P	MILP	£, (CO <sub>2</sub> )	GE, GT, NG-CHP, NGB, AR, ECh, PV, StoreT, StoreT	E, H, C	P: GE & GT	-	-
42	1H-1Y/12TD	12B-AD+N	P	MILP	£, (CO <sub>2</sub> )	NG-CHP, PV, ST, StoreT, NGB	E, H	-	-	Compares heating network routing.
43	1H-3TD	1N-AD	P	MILP	£, (CO <sub>2</sub> )	WT, NG-CHP, NGB, StoreT, StoreE, DR	E, H	-	SO	-
44	1H-40Y/200TD	300B-D	P	MILP (TIMES)	£, (CO <sub>2</sub> )	E/O/W-B, B-CHP, HP, ST, Hydrop, PV	E, H	-	S	-
45	1H-1Y/2TD	1N-AD	P	MILP	£	NGB, CHP, AR, ECh	E, H, C	-	max regret	-
46	1H-15Y/5Y	C	P	MILP	£, (CO <sub>2</sub> , $\eta$ )	CCGT, WT, PV, H2	E	-	-	-
47	4P-25Y/18TD	3N-D	P	MILP	£	CHP, PV + 21 national-scale techs	E	-	RO	Software open-source
48	30min-1D	50B-AD+N	O	MILP	£	CHP, AHP, E-B, StoreT, StoreE	E, H	-	SO	Linearised temperature tracking & thermal comfort considered TDs run in parallel
49	7P-1Y	6B-D+N	P	MILP (CPLEX via GAMS)	£, CO <sub>2</sub>	NGB, AR, ECh, CHP	E, H, C	P: £	RO/SO: £, GridE, CO <sub>2</sub> , demand	-

Table A.3 continued from previous page

Model	Resolution		Optimisation			Model elements				Additional elements
	Temporal	Spatial	Mode	Method	Obj	Technologies	Energy types	NL repr.	Uncertainty analysis	
50	1H-1D	37N-AD+N	O	MILP	£, (losses)	PV, WT	E	-	SO: PV, WT, demand	-
51	1Y-20Y	34B-AD+N	P	MILP	£	WT, PV, StoreE	E	PF	SO: WT, PV, demand, £	Software open-source
52	1H-1D	1N-AD	O	MILP (CPLEX via GAMS)	£	WT, NG-CHP, NGB, StoreT, StoreE, DR	E, H	-	S & SO: demand, £	-
53	1H-1D	37N-AD+N	O	MILP + iteration	£	PV StoreE	E	PF	SO	-
54	1H-1D	26B-D+N	O	2-stage MILP	£	NG-CHP, NGB, PV, AHP, StoreT, StoreE	E, H	PF, HF	RO/SO	-
55	Varying	1N-D	P	MILP	£	ST, PV, StoreE, StoreH2, NG/FC/H2-CHP, NGB, StoreT, AHP	E, H	ML: AHP, GT; P: cost <sup>£</sup>	-	Assesses impact of TDs and inter-cluster storage
56	Varying	1N-4B-AD	P	MILP	£	PV, WT, H2, StoreE, AHP, E-B, StoreT, NG-CHP, NGB, CCGT	E, H	-	-	Assesses impact of TDs and inter-cluster storage
57	1H-1Y/14TD	10B-D+N	P	MILP	£	PV, StoreE, GSHP, NGB, NG-CHP, BM-B, StoreT	E, H	-	SO (inc. risk)	Compares risk aversion techniques
58	1H-1Y	10B-D+N	P	MILP (CPLEX)	£	PV, StoreE, GSHP, NGB, NG-CHP, BM-B, StoreT	E, H	-	SO: £, demand, emission factor	ε-constrained CO <sub>2</sub>
59	1H-1D	1N-AD	O	2-stage MILP	£	PV, WT, B-CHP, StoreE, StoreT, NGB, ICE	E, H	-	S	Heuristic top-level planning
60	1H-1D	15N-AD+N	P	MILP (CPLEX via GAMS)	£	WT, PV, StoreE, ECh, AR, StoreT, NGB, NG-CHP	E, H, C	-	RO: PV, WT, £	Can provide spinning reserve for an external market

# Appendix B

## Case studies

### B.1 Case study 1: Japanese hotel

#### B.1.1 Demand and district network

This case study consists of only one building, a hotel, so there is no network associated with it. The hotel has 20,000m<sup>2</sup> total floor space, with demand for electricity, hot water, and space cooling (Figure B.1) over a 24-hour summer day. However, this allows for greater detail in the internal energy distribution, as given in Figure B.2. Figure B.1 shows gas and electricity prices. Electricity export, available for the [solar photovoltaic panel \(PV\)](#) and [electrical battery storage \(StoreE\)](#) technologies, produces a revenue for the system. The electricity sale price is static throughout the day, while the purchase price varies by dynamic pricing.

**Sources** Dynamic pricing is inferred from the building electricity demand profile alongside a knowledge of the size and electricity price of the network power stations, as formulated by Ikeda and Ooka (2016). Energy demand specific to a hotel was taken from the The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE) (2003) database. Japanese Meteorological Agency data on ambient conditions was used for energy consumption modelling of temperature dependent cooling technologies and available solar energy resource for PVs.

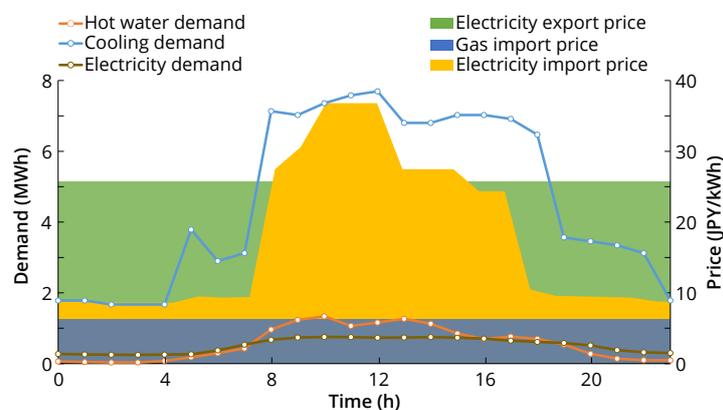


Figure B.1 Case study 1 energy demand and fuel pricing, according to The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE) (2003) and Ikeda and Ooka (2016).

### B.1.2 Available technologies

Technology characteristics are detailed in Table B.1 and are connected together as per the schematic given in Figure B.2. The cooling towers are only used when optimising with the **epsilon constraint differential evolution ( $\epsilon$ DE)** algorithm, where it is possible to track cooling flow temperature.

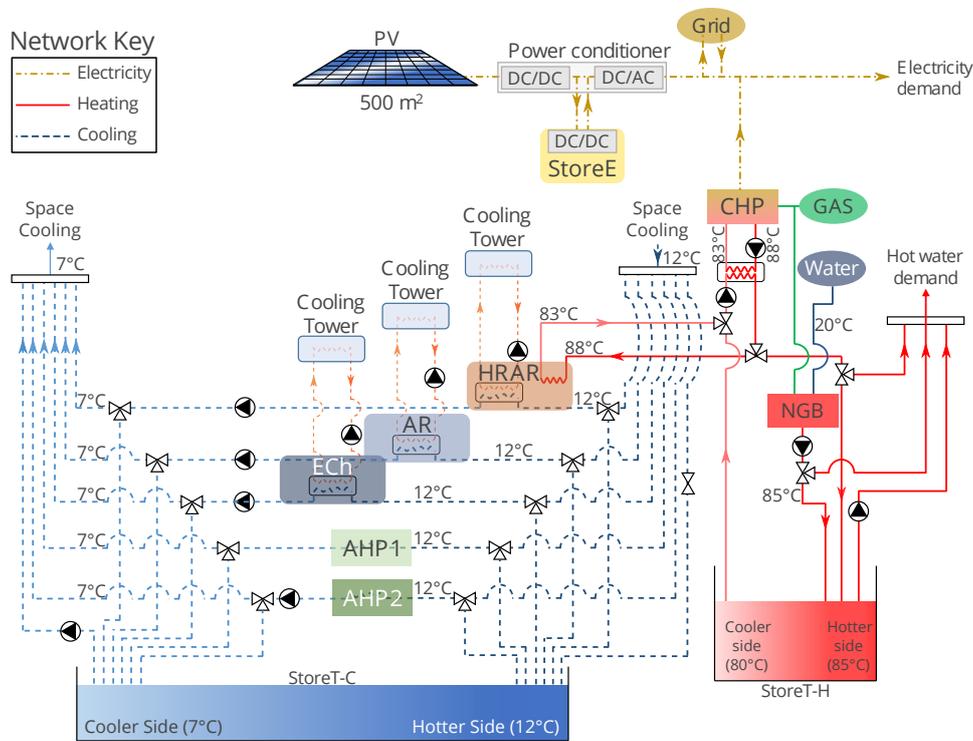


Figure B.2 Case study 1 technology connection schematic.

Table B.1 Case study 1 technology characteristics. E = electricity, G = gas, S = solar radiation, C = cooling, H = heating.

Technology	AHP	ECh	HRAR	AR	CHP <sup>a</sup>	NGB	PV	StoreE	StoreT-H	StoreT-C
Consumption	E	E	G/H	G	G	G	S	E	H	C
Production	C	C	C	C	E, H	H	E	E	H	C
Rated capacity	550kW/500kW <sup>b</sup>	2500kW	1000kW	1500kW	352kWe 300kWth	750kW	1.2kW/m <sup>2</sup>	1200kWh	10000kWh	500kWh
Rated efficiency	358%	603%	132%	125%	40.5% 34.5%	80.0%	85% <sup>c</sup>	90%	100%	100%
Minimum load rate	0.2	0.45	0.30	0.25	0.2	0.2	0	0	0	0
Rated charge/discharge				N/A				360kW	3000kW	100kW

<sup>a</sup> CHP produces both electricity and heat.

<sup>b</sup> Two AHPs, with differing maximum capacity, were available in this study.

<sup>c</sup> Solar energy conversion efficiency has already been accounted for in the timeseries resource data, PV efficiency of 85% refers to inverter efficiency.

**Sources** Technology characteristics are based on The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE) (2003) database guidelines, alongside the combined result of various Japanese manufacturer specifications.

### B.1.3 Optimisation problem

The case study was optimised using a linear solver, IBM ILOG CPLEX 12.6.2 (IBM Corp., 2016), and a custom metaheuristic algorithm, developed by (Ikeda and Ooka, 2016) and optimised using MATLAB (The Mathworks, Inc., 2015). In both models, the objective was to select the technology operation schedule over 24 hours which would minimise total operation costs. Table B.2 gives the timeseries decision variables associated with each technology.

Table B.2 Technologies associated with each decision variable in case study 1. Decision variables are either continuous (C), binary (B), or SOS2 (SOS2). A description of each decision variable is given in Table 3.2 on page 42.

Hotel	Decision Variables						
	E <sup>+</sup> (C)	E <sup>-</sup> (C)	E <sup>-</sup> (SOS2)	E <sup>ex</sup> (C)	U (B)	S (C)	cost <sup>op</sup> (C)
	AHP, ECh, HRAR, AR, CHP, NGB, PV, StoreE, StoreT-H, StoreT-C	NGB, StoreE	AHP, ECh, HRAR, AR, CHP, NGB, StoreE, StoreT-H <sup>a</sup> , StoreT-C <sup>a</sup>	PV, StoreE	AHP, ECh, HRAR, AR, CHP, NGB, StoreT-H <sup>a</sup> , StoreT-C <sup>a</sup>	StoreE, StoreT-H, StoreT-C	AHP, ECh, HRAR, AR, CHP, NGB, PV <sup>b</sup> , StoreE <sup>b</sup> , StoreT-H <sup>a</sup> , StoreT-C <sup>a</sup>

<sup>a</sup> Storage technology pumps are associated with these decision variables.

<sup>b</sup> These technologies export electricity in association with this decision variable.

Table B.3 gives the problem size for the mixed integer linear programming (MILP) model, including the single line efficiency (SLE) and special order set constraint of type 2 (SOS2) piecewise problems. Given that the different optimisation algorithms are used for direct comparison, the same decision variables and constraints exist in the two models. The problem size associated with the same problem, including investment decisions is also given for scale comparison. In the MILP model, binary and integer decision variables associated with the model are for the operation of each technology in each timestep, in order to model a minimum load rate on technologies. Additionally, in the piecewise model, there are SOS2 constraints which account for the increase in the number of binary decision variables. Following preprocessing in the CPLEX solver, a large number of redundant decision variables and constraints are found, leading to a significantly reduced problem, particularly in the SLE problem.

Table B.3 MILP problem size for case study 1, as reported by IBM ILOG CPLEX. \* refers to models which were not considered in the primary analysis, but are shown here for scale. The reduced MIP model refers to the problem which CPLEX ultimately optimises, following a preprocessing step to remove redundant decision variables and constraints (such as those which are set to equal a specific value, so can be deemed as parameters instead of variables.)

Case study	Input Variables			Reduced MIP		
	All	Binary/Integer/SOS2	Constraints	All	Binary/Integer/SOS2	Constraints
SLE	6,683	1,859	8,036	490	214	486
SOS2 piecewise	12,275	2,579	9,860	5,527	896	2,396
SOS2 piecewise, inc. investment decisions*	89,692	15,004	61,225	60,251	10,720	28,496

## B.2 Case study 2: UK district

### B.2.1 Demand and district network

This district is illustrative and consists of 10 domestic properties, one large hotel, one large office, and one power plant. This has been aggregated into the three nodes seen in Figure B.3, in which there is one domestic node (10 domestic properties), one commercial node (hotel and office) and one power plant node. Distribution networks exist for low voltage electricity, gas, and heat.

The district is located in the South-East of England, UK. The demand data, given in Figure B.4 was aggregated to four typical days (96 timesteps) when comparing piecewise to other linearised consumption curves. Two separate weeks (winter & summer, 168 timesteps each) were modelled when analysing equidistant against optimised breakpoint allocation.

Table B.4 gives further information on attributes of each property type, including which technologies are available. No cooling technologies are considered in the domestic properties as there is only electricity and heat demand at those nodes. Roof area is used to constrain the total capacity of solar technologies.

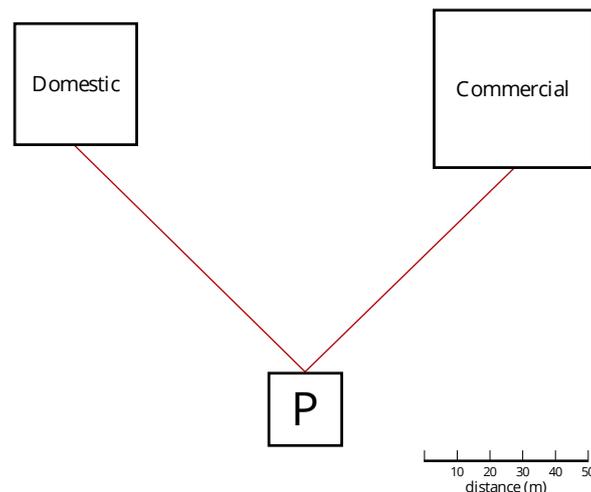
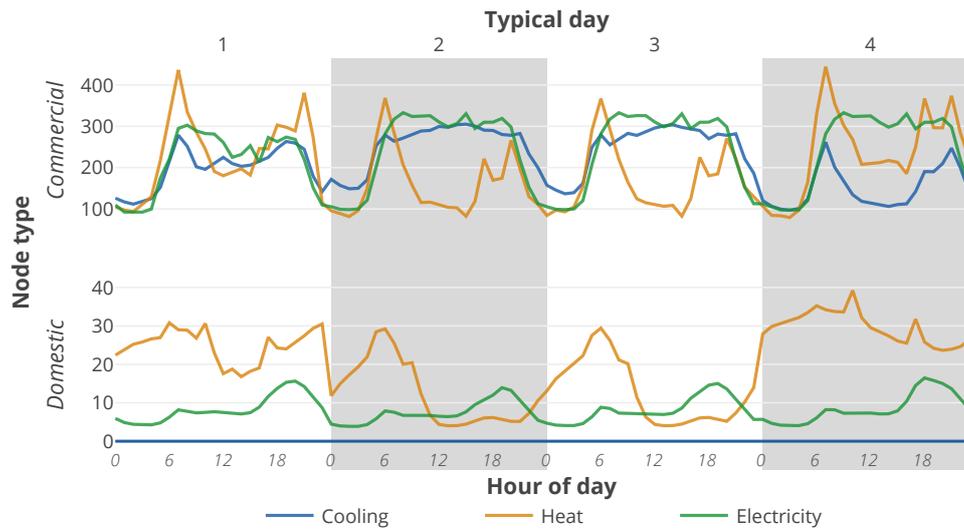


Figure B.3 Graphical representation of case study 2 district network. P = power plant node.

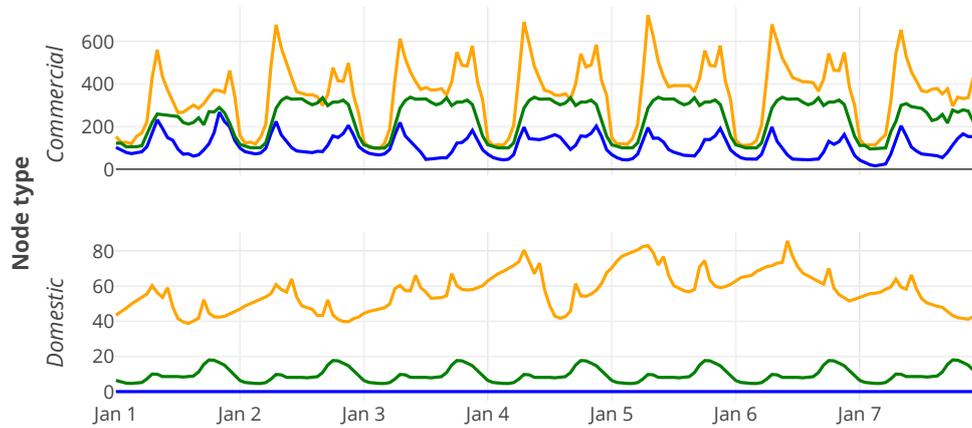
Table B.4 Case study 2 district node details.

	Dwelling	Hotel	Office	Plant
Annual energy demand (MWh)	Electricity	7.2	1595.5	481.3
	Heat	17.5	1641.6	86.5
	Cooling	0.0	1757.9	99.1
Available roof area (m <sup>2</sup> )	130	1300	900	0
Available technologies	NGB, PV, ST, StoreE, StoreT	$\mu$ CHP, NGB, PV, HRAR, AHP, ECh, ST, StoreE, StoreT	NGB, PV, ST, HRAR, AHP, ECh, StoreE, StoreT	CHP, GridE, GridNG

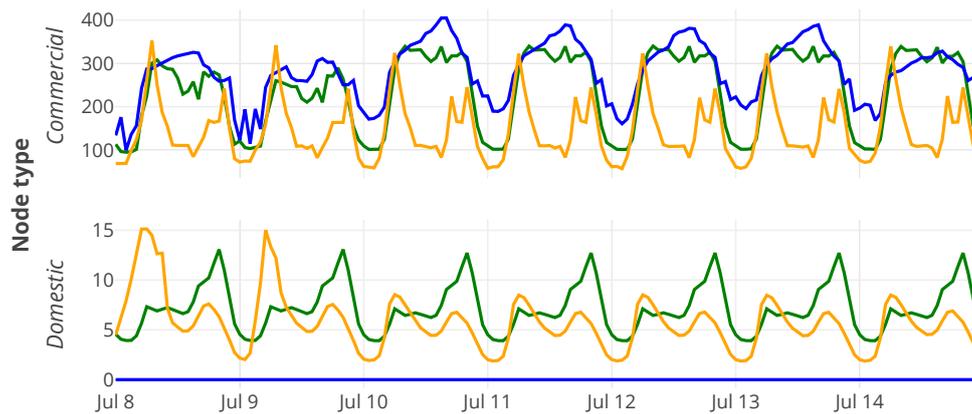
**Sources** U.S. Department of Energy representative building demand (EERE, 2013) informed hourly heat, cooling, and electricity demand of district buildings. Seattle, WA data was chosen due to climate similarity with London, UK.



(a) 4 TDs representing the full timeseries.



(b) Winter week.



(c) Summer week.

— Cooling      — Heat      — Electricity

Figure B.4 Case study 2 input demand. a) shows the timeseries clustered into four TDs; b) and c) shows a winter and summer week, respectively.

## B.2.2 Available technologies

There are a total of 16 technologies available for purchase in the district (10 supply, 3 storage, 3 distribution). The attributes of each technology are given in tables B.5 and B.6. Most technology investment costs are a function of the technology capacity as well as a capacity-independent cost. Operational costs are primarily related to the purchase of **national grid electricity (GridE)** and **national grid natural gas (GridNG)**. Negative costs are also associated with revenue if exporting **combined heat and power plant (CHP)** or **PV** electricity. **PV** and **Solar thermal panel (ST)** available energy is intermittent, based on weather conditions in each hour.

Table B.5 Case study 2 technology characteristics. E = electricity, G = gas, S = solar radiation, C = cooling, H = heating.

Technology	AHP	ECh	HRAR	CHP <sup>a</sup>	NGB	PV	ST	GridE	GridNG	StoreE	StoreT-H	StoreT-C
Consumption	E	E	G/H	G	G	S	S	N/A	N/A	E	H	C
Production	C	C	C	E, H	H	E	H	E	G	E	H	C
Rated capacity	550kW/ 500kW <sup>b</sup>	2500kW	1000kW	352kWe 300kWth	750kW	1.2kW/m <sup>2</sup>	N/A	N/A <sup>c</sup>	N/A <sup>c</sup>	1200kWh	10000kWh	500kWh
Rated efficiency	358%	603%	132%	40.5% 34.5%	80.0%	85% <sup>d</sup>	100% <sup>d</sup>	100%	100%	90%	100%	100%
Minimum load rate	0.2	0.45	0.30	0.2	0.2	0	0	0	0	0	0	0
Rated charge/discharge						N/A				360kW	3000kW	100kW
Investment cost (GBP)	2,517 + 158/kW	111/kW	52,497 + 79/kW	46,480 + 703/kW	2,024 + 35/kW	1,500 + 1,000/kW	1200/kW	75 <sup>e</sup>	16,800 <sup>f</sup>	1,667 + 350/kWh	527 + 66/kWh	527 + 303/kWh
Operation cost (GBP)	0	0	0	0.004/kWh <sup>g</sup>	0	-0.1203/kWh -0.0491/kWh <sup>h</sup>	0	0.095/kWh	0.025/kWh	0	0	0

<sup>a</sup> CHP produces both electricity and heat.

<sup>b</sup> Two AHPs, with differing maximum capacity, were available in this study.

<sup>c</sup> GridE and GridNG capacity is unconstrained.

<sup>d</sup> Solar energy conversion efficiency has already been accounted for in the timeseries resource data, PV efficiency of 85% refers to inverter efficiency.

<sup>e</sup> Fixed charge on per-building tariffs.

<sup>f</sup> Fixed charge to lay 300m of pipeline to the district.

<sup>g</sup> CHP also has possible revenue, at 80% of the wholesale electricity price in any hour.

<sup>h</sup> Revenue gained by the PV feed-in tariff differs between residential (top) and commercial (bottom) properties.

Table B.6 Details of distribution technologies in case study 2. E = electricity, G = gas, H = heating.

Distributed energy	E	G	H
Capacity	2000kW	2000kW	2000kW
Rated efficiency	98%	100%	99.75%/m
Investment cost (GBP)	0 <sup>a</sup>	56/m	530/m

<sup>a</sup> Cost of electrical lines are ignored since electricity lines are assumed necessary, no matter what the method for meeting electricity demand.

**Sources** Supply technology characteristics are based on The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE) (2003) database guidelines. Distribution technology characteristics are based on an aggregation of values in the literature found in Appendix A. Pre-computation of PV available energy was undertaken by Pfenninger and Staffell (2016) and acquired for London from <https://renewables.ninja> (Pfenninger and Staffell, 2016). Using MERRA reanalysis data, also acquired from <https://renewables.ninja>, the heat output

equation proposed by Brunold *et al.* (1994) was used in this case study to pre-compute the ST heat output.

PV feed-in tariff prices are taken at January 2016 value, from ofgem (2016). Hourly GridE export prices are courtesy of ELEXON (<https://www.elexon.co.uk/>). Most costs have been calculated based on values given in the SPON'S mechanical and electrical services price book (AECOM, 2015). These costs include a fixed installation cost and a cost that increases linearly with technology capacity. Thermal energy storage (StoreT), StoreE, PV, and ST costs have been aggregated from online suppliers.

### B.2.3 Optimisation problem

The MILP model was optimised using IBM ILOG CPLEX 12.6.2 (IBM Corp., 2016), with the objective to select technology capacities and their operation schedule of the full time horizon. Table B.8 gives the investment and timeseries decision variables associated with each technology at each location. Binary decision variables associated with the model are for the operation of each technology in each timestep, in order to model a minimum load rate on technologies, and for the purchase of each technology, in order to include a capacity-independent investment cost.

Table B.9 gives the problem size for case study 2, including the SLE and convex bounding sets (CBS) piecewise problems. In the piecewise models, there are initially marginally more constraints, to describe each curve. Following preprocessing in the CPLEX solver, the number of decision variables and constraints reduces by an order of magnitude, leading to a significantly reduced problem, albeit still with several thousand binary decision variables.

Table B.8 Technologies associated with each decision variable in case study 2. Decision variables are either continuous (C), binary (B), or CBS (CBS). A description of each decision variable is given in Table 3.2 on page 42.

		Investment decision variables						
		$\hat{E}$ (C)	$\hat{R}^{area}$ (C)	$\hat{S}$ (C)	$\hat{U}$ (B)	$cost^{inv}$ (C)		
Domestic Plant		CHP, GridE, GridNG			CHP, GridE, GridNG	CHP, GridE, GridNG		
		NGB, PV, ST, StoreE <sup>a</sup> , StoreT <sup>a</sup>	PV, ST	StoreE, StoreT	PV, StoreE, StoreT	NGB, PV, ST, StoreE, StoreT		
Distribution <sup>c</sup> Commercial		$\mu$ CHP, NGB, PV, HRAR, AHP, Ech, ST, StoreE <sup>a</sup> , StoreT <sup>a</sup>	PV, ST	StoreE, StoreT	$\mu$ CHP, PV, HRAR, AHP, StoreE, StoreT	$\mu$ CHP, NGB, PV, HRAR, AHP, Ech, ST, StoreE, StoreT		
		E, G, H			G, H	G, H		
		Timeseries decision variables						
		$E^+$ (C)	$E^-$ (C)	$E^-$ (CBS)	$E^{ex}$ (C)	U (B)	S (C)	$cost^{op}$ (C)
Domestic Plant		CHP, GridE, GridNG	GridE, GridNG	CHP	CHP	CHP		CHP <sup>b</sup> , GridE, GridNG
		NGB, PV, ST, StoreE, StoreT	NGB, StoreE, StoreT		PV	NGB	StoreE, StoreT	PV <sup>b</sup>
Distribution <sup>c</sup> Commercial		$\mu$ CHP, NGB, PV, HRAR, AHP, Ech, ST, StoreE, StoreT	NGB, StoreE, StoreT	$\mu$ CHP, HRAR, AHP, Ech	$\mu$ CHP, PV	NGB, $\mu$ CHP, HRAR, AHP	StoreE, StoreT	$\mu$ CHP <sup>b</sup> , PV <sup>b</sup>
		E, G, H	E, G, H					

<sup>a</sup> Energy capacity = maximum charge/discharge rate for storage technologies.

<sup>b</sup> These technologies export electricity in association with this decision variable.

<sup>c</sup> Distribution technologies are given by the energy they distribute: E = electricity, G = natural gas, H = heat.

Table B.9 MILP problem size for case study 2, as reported by IBM ILOG CPLEX. 1 week models were used to compare the effect of the number of breakpoints; 4 TD models were used to compare CBS piecewise and SLE models. The reduced MIP model refers to the problem which CPLEX ultimately optimises, following a preprocessing step to remove redundant decision variables and constraints (such as those which are set to equal a specific value, so can be deemed as parameters instead of variables.)

Case study	Input Variables		Constraints	Reduced MIP		
	All	Binary/Integer/SOS2		All	Binary/Integer/SOS2	Constraints
1 week, SLE	432,073	6,156	415,034	13,942	3,041	19,248
1 week, 6 breakpoint CBS piecewise	432,073	6,156	417,554	15,794	3,210	21,943
4 TDs, SLE	247,393	3,564	237,699	8,862	1,752	12,465
4 TDs, 3 breakpoint CBS piecewise	247,393	3,564	239,139	9,633	1,848	14,004

## B.3 Case study 3: Bangalore district

This case study is available as a Calliope model at <https://github.com/brynpickering/bangalore-calliope>.

### B.3.1 Demand and district network

A collection of 17 office buildings within Bangalore, India have been selected, defining an illustrative district. Buildings within close geographic proximity have been combined into nodes. Figure B.5 shows the resulting 11 nodes. Most nodes consist of several buildings, which are connected at the same point on the district cooling network. Building floor area (Table B.10) has been inferred from the external footprint and number of floors for each building. As the development is fictitious, no other information is known about these buildings. Total district energy demand is given in Figure B.6, which shows the average, minimum and maximum of the 500 scenarios following clustering of the full year into 12 typical days (TDs). The average demand was used in single-scenario optimisation.

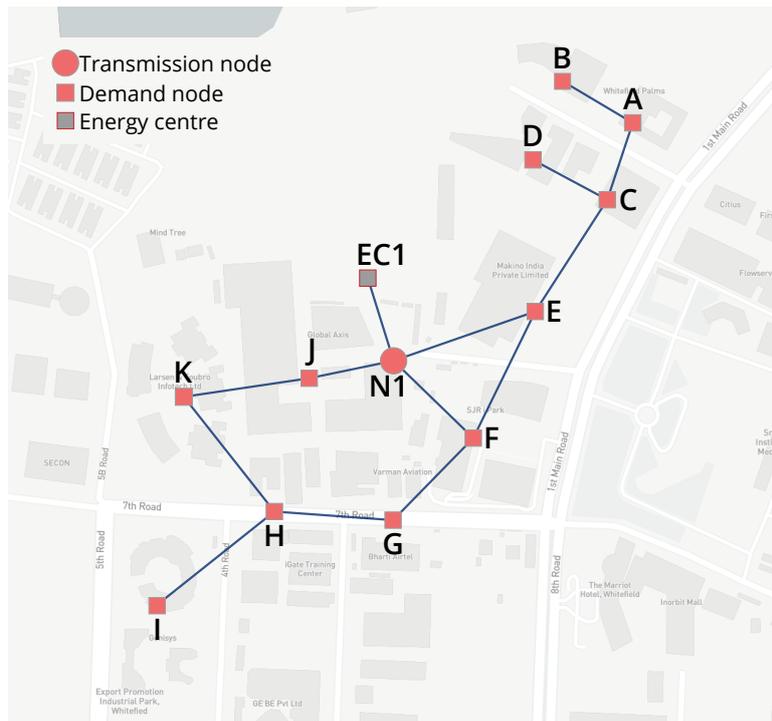


Figure B.5 Case study 3 district network.

Table B.10 Case study 3 district node details. GIA = gross internal floor area. Areas given to three significant figures.

Node	A	B	C	D	E	F	G	H	I	J	K	EC1	N1
GIA ( $\times 10^3 \text{m}^2$ )	5.44	36.6	12.7	22.4	17.2	78.1	46.6	93.1	23.8	178	39.5	N/A	N/A
Roof area ( $\times 10^3 \text{m}^2$ )	2.72	6.10	3.16	5.60	8.59	11.16	11.65	18.61	5.96	22.31	7.90	N/A	N/A
Technologies						PV, ECh, DG, StoreE, GridE						ECh, CCHP, StoreT, StoreE, GridE	

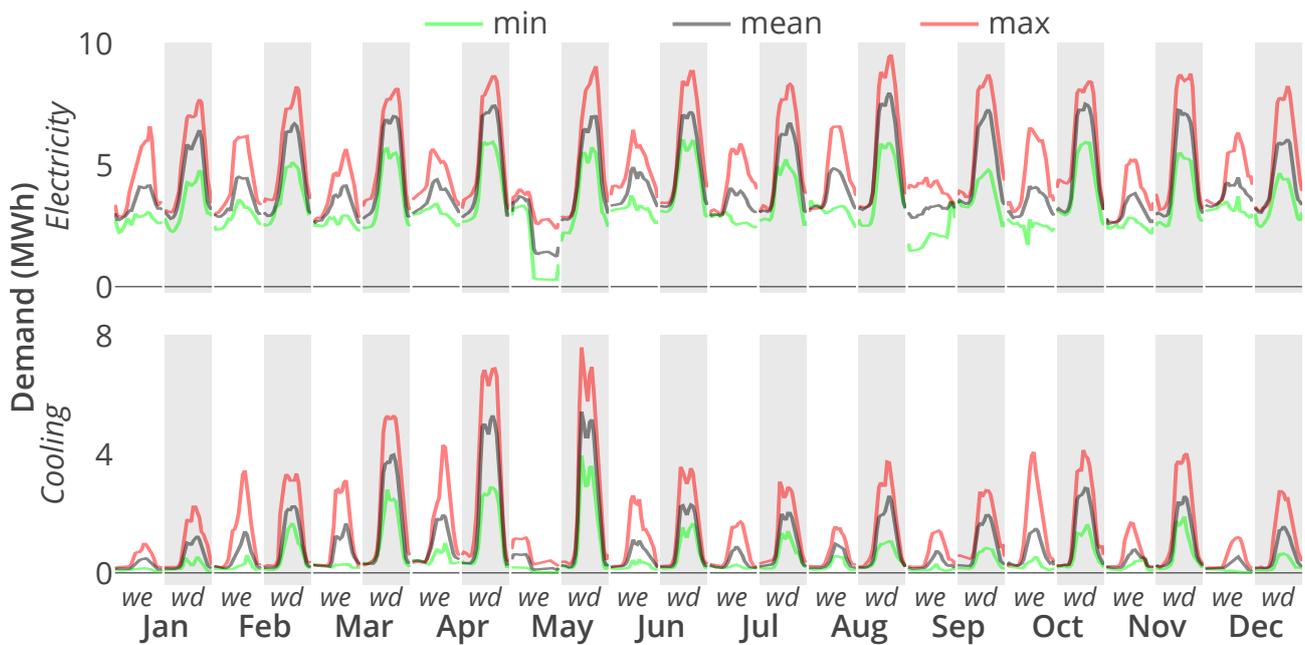


Figure B.6 Typical day demand profiles for case study 3. Three traces are given per typical day. The mean trace shows the average demand across all 500 scenarios and all days within the typical day group. Maximum and minimum traces are the extreme demand profiles from any of the 500 scenarios, on any of the days within the typical day group.

**Sources** Building footprint and heights are taken from OpenStreetMap (OpenStreetMap contributors, 2017) for a subset of buildings in the latitude – longitude bounding box [Top-left: 12.9860 – 77.7210; bottom-right: 12.9765 – 77.7290]. Consumption data was provided by the Indian Institute of Human Settlements for their office in Bangalore. Demand scenarios were generated in Ward *et al.* (2018).

### B.3.2 Available technologies

There are 10 distinct technologies available for purchase in the district (6 supply, 2 storage, 2 distribution), as well as the option to purchase GridE, diesel, and biomass. Table B.10 shows the difference in technologies available for purchase in the energy centre and the demand nodes. Technology characteristics, including purchase and operation costs, are given in Table B.11.

**Sources** Except for the CCHPs, technology costs are based on values provided by an Indian engineering firm. Where costs specific to India were not available, values from the (UK specific) SPON'S mechanical and electrical services price book (AECOM, 2015) have been used, assuming a currency conversion factor of 90:1 INR:GBP.

### B.3.3 Optimisation problem

This model was optimised using IBM ILOG CPLEX 12.6.2 (IBM Corp., 2016) on the CSD3 high performance computing cluster at the University of Cambridge. Table B.12 gives the

Table B.11 Case study 3 technology characteristics. CTP = cooling to power ratio.

	Efficiency or COP	Maximum capacity	Cost <i>Capacity</i>	<i>Operation</i>	Carbon emissions	Other
PV	1	N/A	$5.50 \times 10^4$ INR/kW <sub>p</sub>	0 <sup>b</sup>	0	7m <sup>2</sup> /kW <sub>p</sub>
B-CCHP	0.2	100MW <sub>p</sub>	$4 \times 10^5$ INR + $6.58 \times 10^4$ INR/kW <sup>a</sup>	12 INR/kWh <sup>b</sup>	0.01 kgCO <sub>2</sub> /kWh <sup>c</sup>	CTP: 2.1
D-CCHP	0.45	100MW <sub>p</sub>	$4 \times 10^4$ INR + $2.80 \times 10^5$ INR/kW <sup>a</sup>	23 INR/kWh <sup>b</sup>	0.28 kgCO <sub>2</sub> /kWh <sup>c</sup>	CTP: 0.7
DG	0.45	100kW <sub>p</sub>	$0.90 \times 10^4$ INR/kW	16 INR/kWh	0.28 kgCO <sub>2</sub> /kWh <sup>c</sup>	
ECh	3	1MW <sub>p</sub>	$2.41 \times 10^4$ INR/kW	8 INR/kWh	0.7 kgCO <sub>2</sub> /kWh <sup>c</sup>	
EC-ECh	5	100MW <sub>p</sub>	$2.94 \times 10^4$ INR/kW	8.25 INR/kWh <sup>a</sup>	0.7 kgCO <sub>2</sub> /kWh <sup>c</sup>	
StoreE	1	5MWh	$1.70 \times 10^4$ INR/kWh	0	0	charge rate: 0.7
StoreT	1	100MWh	$3 \times 10^3$ INR/kWh	0	0	charge rate: 0.5
<b>Distribution</b>						
Electricity	1	20MW <sub>p</sub>	0	0		
Thermal	1	20MW <sub>p</sub>	731 INR/kW <sub>p</sub> /m	0		

<sup>a</sup> There is an additional cost of  $5 \times 10^7$  INR for construction of an energy centre, which is incurred when at least one energy centre technology is chosen.

<sup>b</sup> combined cooling, heat and power plants (CCHPs) and PV can export electricity at a rate of -3.40 INR/kWh.

<sup>c</sup> Carbon emissions are given per unit energy consumed.

investment and timeseries decision variables associated with each technology at each node type. Binary decision variables associated with the model are for the purchase of each technology, in order to include a capacity-independent investment cost.

Table B.13 gives the problem size for case study 3, including the single scenario and 10 scenario **scenario optimisation (SO)** models. The objective was to select technology capacities and operation schedules over 12 TDs such that total operation costs are minimised. In SO models, the objective included a weighted sum of operation costs across 10 scenarios and in some cases, a risk aversion component: the **conditional value at risk (CVaR)**. The problem size associated with a single scenario optimised over the full time horizon, instead of using TDs, is given for scale comparison. Binary decision variables associated with the model signal the purchase of technologies which define capacity-independent investment costs (i.e. purchase costs). As with case study 4, the problem size is halved prior to optimisation, following preprocessing in the CPLEX solver.

Table B.12 Technologies associated with each decision variable in case study 3. Decision variables are either continuous (C) or binary (B). A description of each decision variable is given in Table 3.2 on page 42.

Investment decision variables					
	$\hat{E}$ (C)	$\hat{R}^{area}$ (C)	$\hat{S}$ (C)	$\hat{U}$ (B)	$cost^{inv}$ (C)
<b>Energy Centre</b>	ECh, CCHP, StoreT <sup>a</sup> , StoreE <sup>a</sup> , GridE		StoreT, StoreE	CCHP	ECh, CCHP, StoreT, StoreE
<b>Demand node</b>	PV, ECh, DG, StoreE <sup>a</sup> , GridE	PV	StoreE		PV, ECh, DG, StoreE
<b>Distribution<sup>b</sup></b>	E, C			E <sup>c</sup> , C	C
Timeseries decision variables					
	$E^+$ (C)	$E^-$ (C)	$E^{ex}$ (C)	$S^d$ (C)	$cost^{op}$ (C)
<b>Energy Centre</b>	ECh, CCHP, StoreT, StoreE, GridE	ECh, CCHP, StoreT, StoreE	CCHP	StoreT, StoreE	CCHP, GridE
<b>Demand node</b>	PV, ECh, DG, StoreE, GridE	ECh, DG, StoreE	PV	GridE	DG, GridE
<b>Distribution<sup>b</sup></b>	E, C	E, C			

<sup>a</sup> Energy capacity = maximum charge/discharge rate for storage technologies.

<sup>b</sup> Distribution technologies are given by the energy they distribute: E = electricity, G = natural gas, C = cooling.

<sup>c</sup> Electricity binary purchase decision variable is only used at the energy centre connection, to emulate the purchase of the energy centre itself.

<sup>d</sup> There are three additional decision variables associated with storage, to capture storage between **typical days**. See Figure 2.2 on page 14 for more details.

Table B.13 MILP problem size for case study 3, as reported by IBM ILOG CPLEX. \* refers to models which were not considered in the primary analysis, but are shown here for scale. The reduced MIP model refers to the problem which CPLEX ultimately optimises, following a preprocessing step to remove redundant decision variables and constraints (such as those which are set to equal a specific value, so can be deemed as parameters instead of variables.)

Case study	Input			Reduced MIP		
	Variables		Constraints	Variables		Constraints
	All	Binary/Integer		All	Binary/Integer	
12 TDs, single scenario	141,790	6	173,112	71,619	6	74,397
Full timeseries, single scenario*	4,181,968	6	4,761,654	2,046,768	6	1,633,830
12 TDs, 10 scenario SO	1,424,967	6	1,744,468	720,775	6	748,436

## B.4 Case study 4: Cambridge district

This case study is available as a Calliope model at <https://github.com/brynpickering/cambridge-calliope>.

### B.4.1 Demand and district network

Case study 4 is based on intended development by the University of Cambridge. The West Cambridge site is a campus of the University, in which there already exists a number of academic, residential, leisure, and commercial buildings. The plan<sup>1</sup> is to construct 383,000m<sup>2</sup> of new floorspace, through a combination of greenfield and brownfield development (the latter directly replacing current buildings). Of the buildings that are given in the masterplan, those which fit within the archetypes ‘desk-based commercial’, ‘desk-based research’, ‘lab-based commercial’, and ‘lab-based research’ are included in this case study. This leads to a total of 46 buildings, described by 39 nodes in the network shown in Figure B.7. Total district energy demand is given in Figure B.8, which shows the average, minimum, and maximum of the 500 scenarios following clustering of the full year into 6 TDs. The average demand was used in single-scenario optimisation.

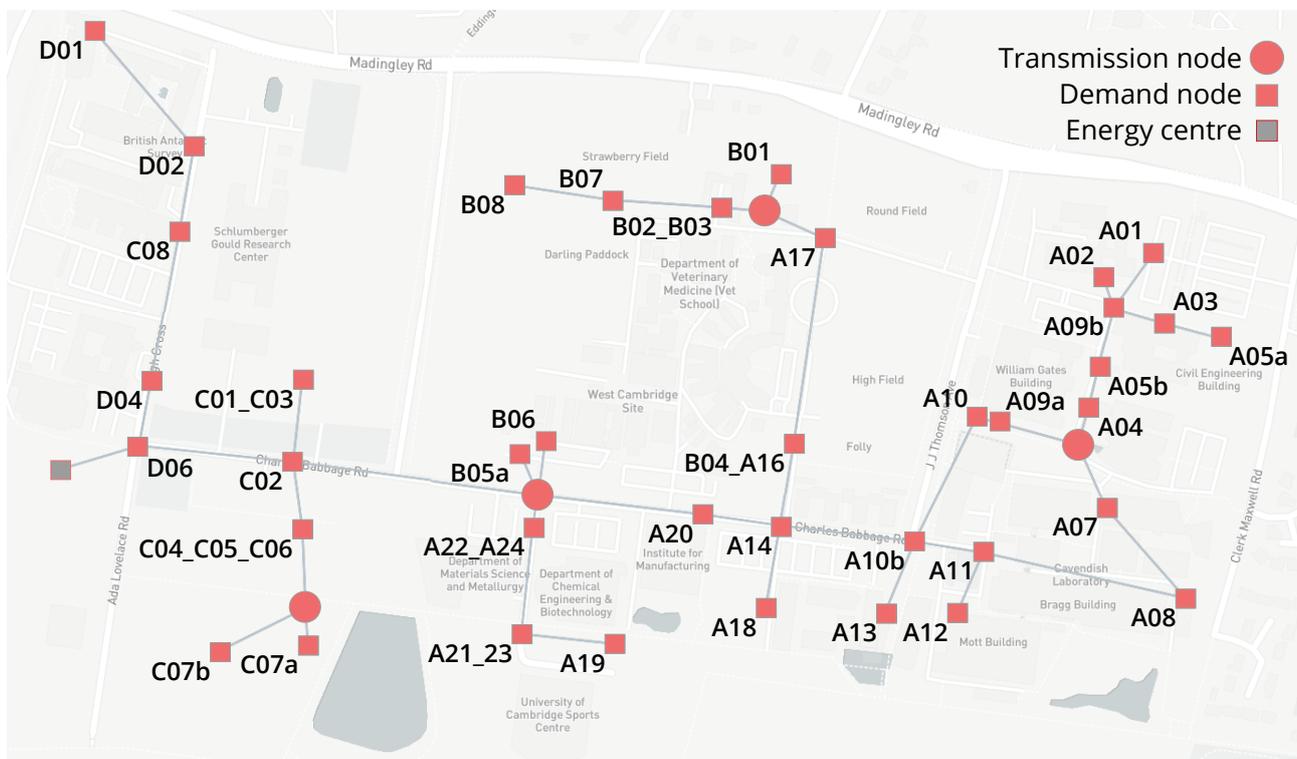


Figure B.7 Case study 4 district network.

<sup>1</sup>More detail on the West Cambridge plan can be found at <http://www.westcambridge.co.uk/>

Table B.14 Case study 4 district node details. GIA = gross internal floor area.

Node	desk_research (10 buildings)	lab_research (17 buildings)	desk_commercial (18 buildings)	lab_commercial (1 building)	Energy centre
GIA	86,774	176,173	181,908	9,473	N/A
Roof area	26,715	72,648	69,824	5,535	
Technologies	PV, NGB, ST, StoreE, StoreT			CHP, GSHP, StoreT, StoreE, GridE, GridNG	

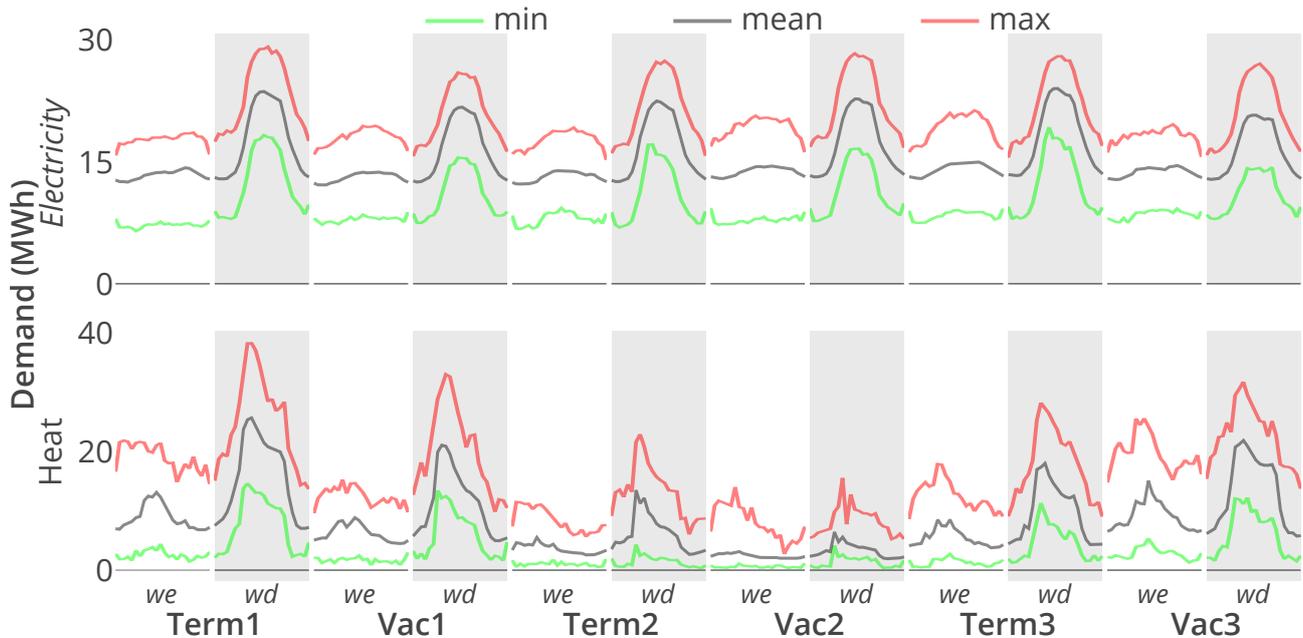


Figure B.8 Typical day demand profiles for case study 4. Three traces are given per typical day. The mean trace shows the average demand across all 500 scenarios and all days within the typical day group. Maximum and minimum traces are the extreme demand profiles from any of the 500 scenarios, on any of the days within the typical day group.

**Sources** Energy consumption data from 17 existing buildings on the University of Cambridge estate were used to inform demand scenario generation. Building archetype categorisation was provided by Aecom, the contracted consultants for the energy plan of the West Cambridge site. Building internal floor area data was provided by the University of Cambridge estates management. Building footprint are taken from OpenStreetMap (OpenStreetMap contributors, 2017).

## B.4.2 Available technologies

There are 10 distinct technologies available for purchase in the district (5 supply, 2 storage, 3 distribution), as well as the option to purchase [GridE](#) and [GridNG](#). Table B.14 shows the difference in technologies available for purchase in the energy centre and the demand nodes. There is no requirement that a given technology is installed at any particular node, as the investment step of the optimisation will decide this. Technology characteristics, including purchase and operation costs, are given in Table B.15.

Table B.15 Case study 4 technology characteristics. HTP = heat to power ratio.

	Efficiency or COP	Capacity	Cost Capacity	Operation	Other
NGB	0.82	2MW <sub>p</sub>	2,024 GBP + 35.3 GBP/kW <sub>p</sub>	0.025 GBP/kWh	
PV	0.85 <sup>a</sup>	N/A	1,500 GBP + 1,000 GBP/kW <sub>p</sub>	-0.1203/kWh -0.0491/kWh <sup>b</sup>	7m <sup>2</sup> /kW <sub>p</sub>
ST	1 <sup>a</sup>	N/A	1,200 GBP + 700 GBP/kW <sub>p</sub>	0	1m <sup>2</sup> /kW <sub>p</sub>
CHP	0.405	25MW <sub>p</sub>	4.65x10 <sup>4</sup> GBP + 703 GBP/kW <sub>p</sub>	0.029 GBP/kWh <sup>c</sup>	HTP: 0.83
GSHP	4.07	22MW <sub>p</sub>	2,520 GBP + 4,221 GBP/kW <sub>p</sub>	0.095 GBP/kWh	
StoreE	1	5MWh	1,670 GBP + 350 GBP/kW <sub>p</sub>	0	charge rate: 0.7
StoreT	0.9	10MWh	527 GBP + 65.7 GBP/kW <sub>p</sub>	0	charge rate: 0.3 heat loss: 0.01
<b>Distribution</b>					
Electricity	1	25MW <sub>p</sub>	0	0	
Heat	2.5x10 <sup>-3</sup> loss/m	25MW <sub>p</sub>	294 GBP + 281 GBP/kW <sub>p</sub> /m	0	
Gas	1	25MW <sub>p</sub>	1.7x10 <sup>4</sup> GBP + 10 GBP/kW <sub>p</sub> /m	0	

<sup>a</sup> Solar energy conversion efficiency has already been accounted for in the timeseries resource data for PV and ST. For PV, 0.85 refers to inverter efficiency.

<sup>b</sup> Revenue gained by the PV feed-in tariff differs between residential (top) and commercial (bottom) properties.

<sup>c</sup> CHP can export to the grid, using a time variable rate equal to 80% of the wholesale electricity prices during 2015.

**Sources** CHP, natural gas boiler (NGB), ground source heat pump (GSHP), and distribution technology investment costs are based on linear interpolation of costs given in AECOM (2015). StoreT, StoreE, PV, and ST costs have been aggregated from online suppliers. Operation costs are taken as average commercial electricity and gas prices in 2015. PV feed-in tariff prices are taken at January 2016 value, from ofgem (2016). CHP export revenue is courtesy of ELEXON (<https://www.elexon.co.uk/>).

CHP and NGB characteristics are based on The Society of Heating Air-Conditioning and Sanitary Engineers of Japan (SHASE) (2003) database guidelines. All other technology characteristics are based on an aggregation of values in the literature found in Appendix A. Pre-computation of PV available energy was undertaken by Pfenninger and Staffell (2016) and acquired for Cambridge from <https://renewables.ninja> (Pfenninger and Staffell, 2016). Using MERRA reanalysis data, also acquired from <https://renewables.ninja>, the heat output equation proposed by Brunold *et al.* (1994) was used in this case study to pre-compute the ST heat output.

### B.4.3 Optimisation problem

This model was optimised using IBM ILOG CPLEX 12.6.2 (IBM Corp., 2016) on the CSD3 high performance computing cluster at the University of Cambridge. Table B.16 gives the investment and timeseries decision variables associated with each technology at each node type. Binary decision variables associated with the model are for the purchase of each technology, in order to include a capacity-independent investment cost.

Table B.16 Technologies associated with each decision variable in case study 4. Decision variables are either continuous (C) or binary (B). A description of each decision variable is given in Table 3.2 on page 42.

Investment decision variables					
	$\hat{E}$ (C)	$\hat{R}^{area}$ (C)	$\hat{S}$ (C)	$\hat{U}$ (B)	$cost^{inv}$ (C)
<b>Energy Centre</b>	CHP, GSHP, StoreT <sup>a</sup> , StoreE <sup>a</sup> , GridE, GridNG	GSHP <sup>b</sup>	StoreT, StoreE	CHP, GSHP, StoreT, GridE, GridNG	CHP, GSHP, StoreT, StoreE, GridE, GridNG
<b>Demand node</b>	PV, NGB, ST, StoreE <sup>a</sup> , StoreT <sup>a</sup>	PV, ST	StoreE, StoreT	PV, NGB, ST, StoreE, StoreT	PV, NGB, ST, StoreE, StoreT
<b>Distribution<sup>c</sup></b>	E, H, G			H, G	H, G
Timeseries decision variables					
	$E^+$ (C)	$E^-$ (C)	$E^{ex}$ (C)	$S^d$ (C)	$cost^{op}$ (C)
<b>Energy Centre</b>	CHP, GSHP, StoreT, StoreE, GridE, GridNG	CHP, GSHP, StoreT, StoreE	CHP	StoreT, StoreE	CHP, GridE, GridNG
<b>Demand node</b>	PV, NGB, ST, StoreE, StoreT	NGB, StoreE, StoreT	PV	StoreE, StoreT	PV
<b>Distribution<sup>c</sup></b>	E, H, G	E, H, G			

<sup>a</sup> Energy capacity = maximum charge/discharge rate for storage technologies.

<sup>b</sup> The GSHP capacity is limited by the extent of the borehole field, hence the need for the  $\hat{R}^{area}$  decision variable.

<sup>c</sup> Distribution technologies are given by the energy they distribute: E = electricity, G = natural gas, C = cooling.

<sup>d</sup> There are three additional decision variables associated with storage, to capture storage between typical days. See Figure 2.2 on page 14 for more details.

Table B.17 gives the problem size for case study 4, including the single scenario and 10 scenario SO models. The objective was to select technology capacities and operation schedules over 6 TDs such that total operation costs are minimised. In SO models, the objective included a weighted sum of operation costs across 10 scenarios and in some cases, a risk aversion component: the CVaR. The problem size associated with a single scenario optimised over the full time horizon, instead of using TDs, is given for scale comparison. Binary decision variables associated with the model signal the purchase of technologies which define capacity-independent investment costs (i.e. purchase costs). This case study provides the largest problem size of all case studies in this thesis. As with case study 4, the problem size is halved prior to optimisation, following preprocessing in the CPLEX solver.

Table B.17 MILP problem size for case study 4, as reported by IBM ILOG CPLEX. \* refers to models which were not considered in the primary analysis, but are shown here for scale. The reduced MIP model refers to the problem which CPLEX ultimately optimises, following a preprocessing step to remove redundant decision variables and constraints (such as those which are set to equal a specific value, so can be deemed as parameters instead of variables.)

Case study	Input			Reduced MIP		
	Variables		Constraints	Binary/Integer		Constraints
	All	Binary/Integer		All	Binary/Integer	
6 TDs, single scenario	229,901	416	400,650	110,196	416	218,024
Full timeseries, single scenario*	11,934,733	416	18,128,034	4,819,588	416	7,201,929
6 TDs, 10 scenario SO	2,887,576	416	3,984,676	1,837,647	416	2,220,271